Investigation of the *cir* multi-gene family of *Plasmodium chabaudi*

Jennifer Catherine Lawton

August 2011

Division of Parasitology
MRC National Institute for Medical Research
The Ridgeway
Mill Hill, London
NW7 1AA

Division of Infection and Immunity
University College London

This thesis is submitted to University College London for the degree of Doctor of Philosophy

Declaration

I, Jennifer Catherine Lawton confirm that the work presented in this thesis is my own. Information derived from other sources has been indicated as such and collaborations have been described in the acknowledgements.

Acknowledgements

I am very grateful to Jean Langhorne for giving me the opportunity to do this project and the freedom to try so many approaches, her supervision and scientific enthusiasm.

The Langhorne lab: My collaborator on this project, Thibaut Brugat, who read through several chapters of this thesis, and worked in collaboration to produce Figure 37, Figure 38 and Figure 39, shown in chapter 5. Deirdre Cunningham, who helped me to design the synthetic *cir* genes and RT-qPCR experiments, provided a huge amount of support for countless techniques throughout my PhD and read through several chapters of this thesis. Bill Jarra who supervised my animal training and read through chapter 5. Anna Sponaas who helped me to design the flow cytometry experiment. Dorothy Ng helped me to design ELISA experiments, and read chapter 1. Ana Paula Rosario also read this chapter and helped me to collect sera, with Christine Tshitenge. Elizabeth Couper for technical assistance in cloning the synthetic *cir* genes PCHAS_070130 and PCHAS_040110, and their electroporation into *P. pastoris.* The rest of the lab for their support and advice, encouragement and all the cups of tea.

NIMR: My thesis advisory committee Tony Holder and Mike Blackman. Chrislaine Withers-Martinez, who gave me advice about protein expression, DLS, coordinated the crystallization trial for PCHAS_000100 and read through chapter 4. CD and aromatic absorbance spectra were carried out by Steve Martin, division of physical biochemistry. The crystallization trial was carried out by Lesley Haire, division of molecular structure. All the support staff in NIMR, especially in LLY for looking after my mice.

WTSI: Annotation of *cir* genes was carried out under the supervision of Ulrike Boehme and Arnab Pain. Recombination and phylogenetic analyses were advised by Andrew Jackson and Simon Harris. Identification of divergent *cir* genes was performed using HMM analyses of *P. chabaudi* ORFs, by the PFAM group. RNA sequencing was carried out in collaboration with Thomas Otto, Adam Reid, Arnab Pain and the technical assistance of the Illumina sequencing team.

I would like to say a huge thank you to my parents for their encouragement and unfailing support, and the rest of my family and friends. Finally, to Dave, who kept me going throughout the PhD, fed me after all the long hours in the lab and helped me to laugh at myself no matter what …*thanks, you were ruddy brilliant!*

Contents

Abstract

The *pir* genes comprise the largest multi-gene family in *Plasmodium*, with members found in *P. vivax*, *P. knowlesi* and rodent malaria species. Despite their almost universal presence, little is known about the functions of the PIR proteins. To investigate the role PIR proteins play during the erythrocytic stages of infection, the *P. chabaudi* model was chosen, where this gene family is termed *cir*. 198 *cir* genes were identified in the *P. chabaudi* genome, 86% of which clustered to form two major sub-families on the basis of sequence similarity.

Quantitative RT-PCR and Illumina RNA sequencing were used to investigate *cir* transcription during *P. chabaudi* infection. Both methods detected many *cir* genes transcribed at low levels, as shown previously for other *pir*s.

Three of the transcribed *cir* genes were selected for recombinant protein expression in the yeast *Pichia pastoris*: PCHAS_000100, PCHAS_070130 and PCHAS_040110. Soluble PCHAS_000100 was used for measurements of CIR secondary structure. Conserved and sub-family specific peptides were also synthesized. Antibodies present in the sera of *P. chabaudi* immune mice recognized all CIR proteins and peptides.

Polyclonal antibodies were used to determine CIR localization by confocal microscopy and flow cytometry. Whilst most CIRs were located within the parasites, some CIRs were also found on the infected erythrocyte surface of mature trophozoites. In addition, CIRs were detected at the apical end of merozoites.

These results imply that CIR proteins are exposed to the immune system during *P. chabaudi* infection and are antigenic, yet immunization with most CIR proteins and peptides did not protect mice from *P. chabaudi* infection. Upon *P. chabaudi* challenge of mice immunized with CIR sub-family specific reagents, increased levels of *cir* transcripts belonging to the other major sub-family were detected. This may explain why few differences in parasitaemia were observed. The exception was observed during *P. chabaudi* challenge of mice immunized with PCHAS_000100, which were able to clear parasitaemia earlier than controls.

List of figures

List of tables

Abbreviations

3’	Three prime
5’	Five prime
3D7	*Plasmodium falciparum* isolate
α	Alpha
Ab	Antibody
AHNAK	Neuroblast differentiation-associated protein
A/J	*Mus musculus* strain
aLRT	Approximate likelihood ratio test
AMA1	Apical merozoite protein 1
ANKA	*Plasmodium berghei* clone ANKA
APC	Antigen presenting cell
AS	*Plasmodium chabaudi* clone AS
β	Beta
BALB/c	*Mus musculus* strain
BIONJ	Algorithm based upon the neighbour joining method
BIR	*Plasmodium berghei* interspersed repeat
BLAST	Basic local alignment tool
βME	Beta-mercaptoethanol
bp	Base pair
BSA	Bovine serum albumin
C3H	*Mus musculus* strain
C57Bl/6	*Mus musculus* strain
C-terminal	Carboxy terminal
CD	Cluster of differentiation or Circular dichroism
CD-1	*Mus musculus* strain
cDNA	Complementary deoxyribonucleic acid
CIDR	Cysteine interdomain rich
CIR	*Plasmodium chabaudi* interspersed repeat
CO_2	Carbon dioxide
CpG	Cytosine-phosphate-guanosine
Cq	Quantification cycle
CSA	Chondroitin sulphate A
CSP	Circumsporozoite protein

CSS	Conservation shifting sites
σ / Δ	Delta
DBL	Duffy binding like
DC	Dendritic cell
DD2	*Plasmodium falciparum* isolate
DDT	Dichlorodiphenyltrichloroethane
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleic triphosphate
ε	Epsilon
EDTA	Ethylene-diamine-tetra-acetic acid
ELISA	Enzyme linked immunosorbent assay
EXP2	Exported protein 2
Fc	Fragment crystallisable portion of immunoglobulin
FCS	Foetal calf serum
FITC	Fluorescein
FSc	Forward scatter
γ	Gamma
g	Gram
x *g*	Relative centrifugal force
γσ T	Gamma delta T cell
gDNA	Genomic deoxyribonucleic acid
GFP	Green fluorescent protein
GPI	Glycosylphosphattidylinositol
HA	Hyaluronic acid
HB3	*Plasmodium falciparum* isolate
HCl	Hydrochloric acid
HEK	Human embryonic kidney
HMM	Hidden Markov model
HSP	Heat shock protein
ICAM1	Inter-Cellular Adhesion Molecule 1
IFNγ	Interferon gamma
Ig	Immunoglobulin
IgA	Immunoglobulin A
IgD	Immunoglobulin D

IgE	Immunoglobulin E
IgG	Immunoglobulin G
IgM	Immunoglobulin M
IL	Interleukin
IL7Rα	Interleukin 7 receptor alpha
i/p	Intra-peritoneal
iRBC	Infected red blood cell
IT4	*Plasmodium falciparum* isolate
i/v	Intra-venous
KCl	Potassium chloride
KAHRP	Knob associated hitidine rich protein
KIR	*Plasmodium knowlesi* interspersed repeat
KLH	Keyhole limpet cyanin
λ	Lambda
LB	Luria Bertani
LG	Le Gascuel
LRT	Likelihood ratio test
μ	Mu
M	Molar
MACS	Magnetic cell separation
MC	Maurer's clefts
MEGA	Molecular Evolutionary Genetics Analysis
MEME	Multiple EM for motif elicitation
mg	Milligram
$MgCl_2$	Magnesium chloride
ML	Maximum likelihood
mM	Millimolar
μl	Microlitre
mRNA	Messenger ribonucleic acid
MSP1	Merozoite surface protein 1
MSP2	Merozoite surface protein 2
MSP8	Merozoite surface protein 8
Muscle	Multiple sequence alignment by log- expectation algorithm
N-terminal	Amino-terminal
NaCl	Sodium chloride

NaOH	Sodium hydroxide
ng	Nanogram
Ni^{2+}-NTA	Nickel nitrilotriacetic acid
NIMR	National Institute for Medical Research
NJ	Neighbour joining
NK	Natural killer cell
NK-T	Natural killer T cell
NNI	Nearest neighbour interchanges
NTC	No template control
OD	Optical density
ORF	Open reading frame
PAM	Pregnancy associated malaria
PAMP	Pathogen associated molecular pattern
PBMC	Peripheral blood mononuclear cell
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PECAM1	Platelet endothelial cell adhesion molecule
PEXEL	*Plasmodium* export element
PFAM	Protein family
*Pf*EMP1	*Plasmodium falciparum* erythrocyte membrane protein 1
*Pf*MC-2TM	*Plasmodium falciparum* Maurer's cleft - two transmembrane domain
PHI	Pairwise homoplasy index
PIR	*Plasmodium* interspersed repeat
pM	Pico molar
PNEP	PEXEL-negative exported protein
p-NPP	P-nitrophenyl phosphate
pPIC9k	Vector used for *Pichia pastoris* electroporation
PRR	Pathogen recognition receptor
PTEX	*Plasmodium* translocon for protein export
PV	Parasitophorous vacuole
qPCR	Quantitative polymerase chain reaction
RAG2	Recombinase activating gene 2
$RAG2^{-/-}$	*Mus musculus* strain containing a targeted disruption in recombinase activating gene 2
RBC	Red blood cell

rif/RIFIN	Repetitive interspersed repeat
RNA	Ribonucleic acid
Rpm	Revolutions per minute
rRNA	Ribosomal ribonucleic acid
RSS	Rate shifting sites
RT	Reverse transcription
SBP1	Skeleton binding protein 1
SD	Standard deviation
SDS-PAGE	Sodium dodecyl-sulphate polyacrylamide gel electrophoresis
SEM	Standard error of the mean
SH-like	Shimodaira-Hasegawa-like
SICA	Schizont-infected cell agglutination
SNP	Single nucleotide polymorphism
SSAHA	Sequence Search and Alignment by Hashing Algorithm
SSc	Side scatter
STEVOR	Sub-telomeric variable open reading frame
surf/SURFIN	Surface-associated interspersed genes
TCR	T cell receptor
TEMED	Tetramethylethylenediamine
Th	T helper cell
TM	Transmembrane
TNFα	Tumor necrosis factor alpha
TOPALi	Tree topology-related analysis of alignments interface
tRNA	Transfer ribonucleic acid
UPGMA	Unweighted pair group method with arithmetic mean
Ups	Upstream
v	Volts
VAR2CSA	Conserved *Pf*EMP1 variant
VCAM1	Vascular cell adhesion molecule 1
VIR	*Plasmodium vivax* interspersed repeat
VSA	Variant surface antigen
VTS	Vacuolar transport signal
WHO	World Health Organization
w/v	Weight in volume
YIR	*Plasmodium yoelii* interspersed repeat

List of buffers and solutions

Coomassie blue staining solution

0.25% w/v coomassie brilliant blue R250, 45% ethanol, 10% acetic acid.

Coomassie de-staining solution

45% ethanol, 10% acetic acid.

Electroporation solution

270 mM sucrose, 10 mM Tris-HCl, 1 mM $MgCl_2$, pH 7.5

ELISA blocking buffer

PBS containing 1% BSA, 0.3% Tween 20 and 0.05% Sodium azide

ELISA wash buffer

PBS containing 0.9% sodium chloride, and 10% Tween 20, pH 7.2.

ELISA diethanolamine buffer

0.26 M diethanolamine, 8 mM magnesium chloride, pH 9.8

Flow cytometry buffer

PBS containing 1% BSA (w/v), 2 mM EDTA and 0.01% sodium azide.

Giemsa stain buffer

20mM Na_2HPO_4, 40mM KH_2PO_4, pH 7.4

Krebs saline

114 mM NaCl, 4.57 mM KCl, 1.15 mM $MgSO_4$ (Krebs and Eggleston, 1940)

MACS buffer

2 mM EDTA and 2% foetal calf serum in PBS

Ni^{2+} NTA binding buffer

50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole, pH 8.0

Ni^{2+} NTA wash buffer

50 mM NaH_2PO_4, 300 mM NaCl, 20 mM imidazole, pH 8.0

Ni^{2+} NTA elution buffer

50mM NaH_2PO_4, 300 mM NaCl, 200 mM imidazole (or 250 mM, 500 mM, 1 M imidazole), pH 8.0

SDS-PAGE resolving buffer

1.5 M Tris base (Sigma), containing 0.4% SDS, pH 8.8

SDS-PAGE stacking buffer

0.5 M Tris base (Sigma), containing 0.4% SDS, pH 6.8

Western blot transfer buffer

Containing 10% methanol, 0.025 M Tris base and 0.2 M Glycine

Western blot blocking buffer

PBS containing 0.1% Tween 20 and 3% bovine serum albumin

Western blot wash buffer

PBS containing 0.05% Tween 20

List of media

BMGY medium

Liquid broth containing 1% yeast extract, 2% peptone, 100 mM potassium phosphate pH 6.0, 1.34 % yeast nitrogen base, $4x10^{-5}$ % biotin and 0.5 % glycerol

BMMY medium

Liquid broth containing 1% yeast extract, 2% peptone, 100 mM potassium phosphate pH 6.0, 1.34 % yeast nitrogen base, $4x10^{-5}$ % biotin and 0.5 % methanol

LB broth

1% tryptone, 0.5% yeast extract, 1% NaCl, pH 7.0

Minimal methanol medium

Liquid broth containing 1.34% yeast nitrogen base, $4x10^{-5}$% biotin and 0.5% methanol

RPMI 1640 (Gibco) complete medium

To which 6 mM HEPES, 0.5 mM sodium pyruvate, 2 mM L-glutamine, 50 µM β-Mercapto-ethanol (Invitrogen) and 10% heat inactivated foetal calf serum (PAA laboratories, Pasching, Austria) were added

RPMI 1640 (Gibco) incomplete medium

To which 6 mM HEPES and 2 mM L-glutamine (both Invitrogen) were added

Yeast extract peptone medium

Containing 1% yeast extract, 2% peptone and 2% dextrose

List of antibodies used

Capture goat anti-mouse Ig (Southern Biotechnology) used for ELISA: 2 μg/ml

Primary Abs:

Ter119 conjugated to biotin (Biolegend)

- Recognizes a *M. musculus* Glycophorin A associated protein (Kina et al., 2000)
- Western blot: 32 pg/ml
- Immuno-fluorescence: 50 μg/ml

NIMP23

- Recognizes *P. chabaudi* $MSP1_{p21}$ (Kina et al., 2000, Boyle et al., 1982, McKean et al., 1993)
- Western blot: 18 μg/ml
- Immuno-fluorescence: 50 μg/ml

Anti-histidine tag (Novagen)

- Western blot: 0.5 μg/ml

Purified mouse IgM, IgG1, IgG2a, IgG2b, and IgG3 (Sigma-Aldrich)

used for ELISA: 2 μg/ml

Polyclonal Anti-CIR Abs generated in mice

- Flow cytometry: diluted 1/50

Polyclonal Anti-CIR Abs generated in rabbit

- Western blot: diluted 1/500
- ELISA: diluted 1/10,000
- Immuno-fluorescence: diluted 1/50

Secondary Abs for ELISA:

- Goat anti-mouse IgM (Southern Biotechnology): 2 μg/ml
- Goat anti-mouse IgG1 (Southern Biotechnology): 2 μg/ml
- Goat anti-mouse IgG2a (Southern Biotechnology): 2 μg/ml

- Goat anti-mouse IgG2b (Southern Biotechnology): 2 µg/ml
- Goat anti-mouse IgG3 (Southern Biotechnology): 2 µg/ml

Secondary reagents for flow cytometry:

- Goat anti-mouse IgG conjugated to FITC (Southern Biotechnology): 10 µg/ml
- Hoechst 33342 (Invitrogen): 10 µg/ml

Secondary reagents for immuno-fluorescence:

- Goat anti-rabbit IgG conjugated to Alexa 680 (Molecular Probes): 10 µg/ml
- Goat anti-mouse IgG conjugated to Alexa 594 (Molecular Probes): 10 µg/ml
- Streptavidin conjugated to Alexa 750-Allophycocyanin (Molecular Probes): 10 µg/ml
- 4',6-Diamidino-2-phenylindole (DAPI, Sigma Aldrich): 1 mg/ml

Secondary Abs for Western blot:

- Horseradish peroxidase (HRP)-conjugated goat anti-mouse IgG, (Southern Biotechnology): 1:10,000 dilution
- Horseradish peroxidase (HRP)-conjugated Abs: goat anti-rabbit IgG, goat anti-mouse IgG and goat anti-rat igG (Biorad): 1:10,000 dilution

Chapter 1: Introduction

1.1 Malaria is a global burden

Half of the world's population is at risk of malaria infection and the disease results in over a million deaths every year (WHO, 2008). The incidence is particularly high in children under five years of age, with an estimated 0.8% infant mortality in Africa alone (Rowe et al., 2006).

Malaria is caused by *Plasmodium* parasites, of which four have been historically known to infect humans: *P. falciparum*, *P. vivax*, *P. ovale* and *P. malariae*. Infection is characterized by relapsing fever, however there are severe syndromes that can also manifest. *P. falciparum* is widely attributed as the most deadly species, because these infections have a high propensity to develop into the severe syndromes listed in Table 1. The other *Plasmodium* species have been considered more benign, however, reports have emerged in recent years not only that *P. vivax* is capable of causing severe disease [reviewed by (Price et al., 2009)], but also a fifth human malaria exists: *P. knowlesi* (White, 2008).

Table 1: Severe malaria disease syndromes.

Syndromes	Clinical features	Disease mechanisms	References
Severe anaemia	Shock, Respiratory distress	Increased erythrocyte destruction (phagocytosis, parasite- and complement- lysis) Reduced erythrocyte production	(Menendez et al., 2000, Phillips et al., 1986, Stoute et al., 2003, Devine, 1991, Sexton et al., 2004)
Cerebral malaria (CM)	Impaired consciousness Convulsions, Long term neurological deficits	Microvascular obstruction (parasites, platelets, rosettes, microparticles) Pro-inflammatory cytokines, Parasite toxins (eg glycosylphosphatidylinositol)	(Taylor et al., 2004, Grau et al., 1987, Xiao et al., 1999, MacPherson et al., 1985, Newbold et al., 1997, Pain et al., 2001, Grau et al., 2003, Wassmer et al., 2004, Allen et al., 1999, Rowe et al., 1995, Cockburn et al., 2004), (Rogerson et al., 1999)[1]
Metabolic acidosis	Respiratory distress, Hypoxia, Acidaemia, Reduced CVP.	Reduced tissue perfusion (eg hypovolaemia), Parasite products, Pro-inflammatory cytokines, Pulmonary pathology (eg airway obstruction)	(Marsh et al., 1995, English et al., 1996, Clark and Cowden, 2003), (Grau et al., 1989)[2], (Kwiatkowski, 1990)[3], (Day et al., 1999)[4],
Other	Hypoglycaemia, Disseminated intra vascular coagulation	Parasite products and/or toxins, Pro-inflammatory cytokines, Cytoadherance	(Clark and Cowden, 2003, Grau et al., 1989, Luty et al., 1999, Anstey et al., 1996, Clark et al., 2003) [1 2 3 4]

Adapted from (Mackintosh et al., 2004). Central venous pressure was abbreviated to CVP.

It appears that in South East Asia, *P. vivax* infection is capable of progressing to severe disease as often as *P. falciparum*, with a similar proportion of fatal infections (Genton et al., 2008, Kochar et al., 2005, Tjitra et al., 2008, Anstey et al., 2009). In addition, *P. vivax* infection during pregnancy can cause miscarriage or decreased full term birth weight (Nosten et al., 1999). *P. knowlesi* is classed as an 'emerging human pathogen', transmitted zoonotically from its natural hosts: the long- and pig-tailed macaques (*Macaca fasicularis and Macaca nemestrina*). In fact, almost a third of malaria cases in Sarawak, Malaysia, were due to *P. knowlesi* infection, rising to 77% in some regions (Cox-Singh et al., 2008).

The repeated nature of infection means that malaria exerts an enormous socio-economic cost to endemic countries - estimates suggest that *P. falciparum* malaria may have reduced the Gross Domestic Product of some African countries by almost 10% (Gallup and Sachs, 2001), and *P. vivax* malaria could have a global burden of US $1.4 - 4.0 billion per year (Price et al., 2009).

In developed countries, malaria has largely been eradicated by widespread drainage and the use of toxic insecticides such as DDT, which purged the vector population: female *Anopheles* mosquitoes, in the mid 20th century. However, such insecticides are now known to be environmentally hazardous and mosquitoes have developed resistance to those still in common usage, thus alternative approaches are sought to control the prevalence of this disease in the 99 countries where malaria remains endemic (Feachem et al., 2010).

Not only have malaria vectors developed resistance to insecticides, but increased drug resistance of parasites is also a growing problem. Resistance to chloroquine first emerged in East Africa in 1978 (Trape, 2001), a phenomenon now found worldwide and foci of artemisinin resistant parasites have been detected recently on the Cambodia-Thailand border (Yeung et al., 2009). This is particularly worrying since artemisinin based combination therapy is the frontline defence for patients with severe *P. falciparum* malaria (WHO, 2008). Clearly, multiple and coordinated approaches will be required for future malaria control in the face of such drug resistance, even if the desired 'elimination era' is not attainable, particularly in Africa (Moonen et al., 2010, Tatem et al., 2010). A major advantage in this struggle would be the development of a malaria vaccine, which ideally would not only protect individuals from infection, or at least from severe disease, but would also reduce the possibility of parasite transmission and therefore re-infection.

1.2 *Plasmodium*

1.2.1 The parasite life-cycle

Plasmodium parasites have a complex life cycle, involving two hosts and development into several morphologically distinct forms, summarized in Figure 1.

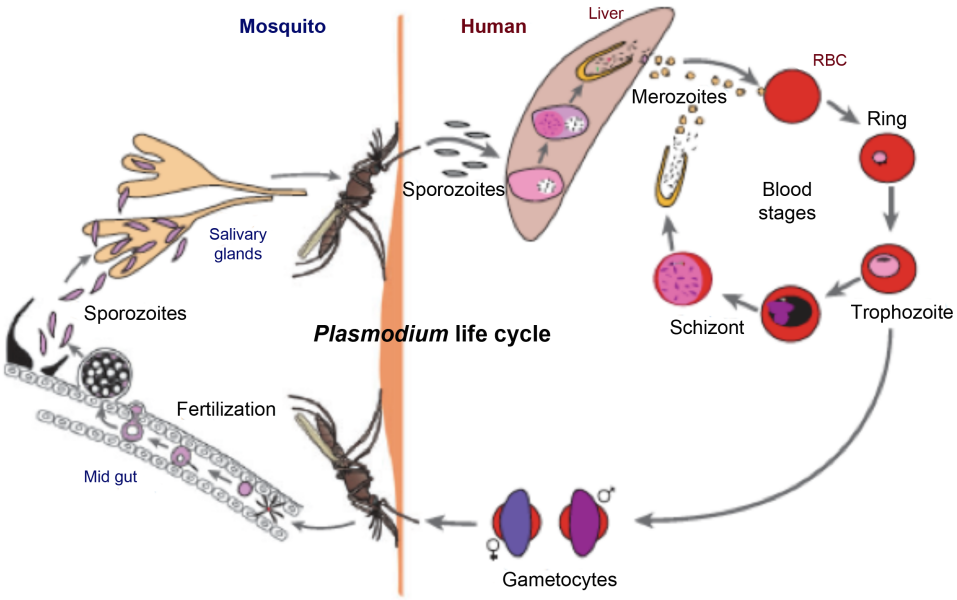


Figure 1: The *Plasmodium* life-cycle.

Adapted from: http://www.cdc.gov/malaria/biology/life_cycle.htm. Sporozoites injected from an infected *Anopheles* mosquito migrate to the liver, where they invade hepatocytes and merozoites develop. Upon release from the liver, merozoites enter the bloodstream where they undergo asexual multiplication in the erythrocytes. After infecting red blood cells, parasites mature into ring stages, trophozoites and then into schizonts, which rupture releasing further merozoites. During the erythrocytic replicative stages, some parasites also differentiate into gametocytes, which undergo sexual maturation when taken up by an *Anopheles* host. Within the mosquito, sexual reproduction occurs, producing a motile zygote known as an ookinete. Following penetration of the gut wall, ookinetes differentiate into oocysts, which, rupture on maturity to release sprorozoites. The sporozoites migrate to the mosquito salivary glands, where they may be injected into another mammalian host.

The definitive host, where sexual reproduction takes place, is the female *Anopheles* mosquito, which injects infective sporozoites into the skin of the second, mammalian host during a blood meal. In the second host, sporozoites migrate through the skin for several hours (Yamauchi et al., 2007), with up to 80% reaching the blood vasculature and completing their migration to the liver (Frevert et al., 2005), where they invade hepatocytes. *Plasmodium* development within a single hepatocyte results in the production of thousands of merozoites, which are released into the blood stream.

Recently, it has been discovered that, at least in the rodent model *P. yoelii*, merozoites are released into the bloodstream within a protective membrane, the merozome, which protects the new merozoites from liver resident phagocytes. Merozomes were observed to sequester in the lung, where they ruptured to release the merozoites (Baer et al., 2007).

Certain species such as *P. vivax* are able to form quiescent uni-nucleate hyponozoites after hepatocyte invasion, such that the parasites may re-enter the cell cycle and give rise to infective parasites several months after initial sporozoite injection (Krotoski et al., 1982). This adaptation likely facilitates *Plasmodium* transmission in areas where the dry season prevents continual breeding of the mosquito vector.

Once released into the blood stream, merozoites begin red blood cell (RBC) invasion, involving attachment of parasites to the RBC surface, followed by re-orientation and invagination of the RBC membrane to form a parasitophorous vacuole [PV; reviewed by (Cowman and Crabb, 2006)]. Within the PV parasites develop first into ring stages, then trophozoites, and finally multi-nucleate schizonts. Rupture of schizonts releases more merozoites to continue cycling through this asexual development process. During this time, some parasites will begin to differentiate into pre-sexual stages: the male and female gametocytes. When these are taken up in the blood meal of an *Anopheles* mosquito, gametocytes mature into reproductively capable gametes and sexual reproduction may occur. This process eventually results in the generation of further infective sporozoites for continuation of the parasite life-cycle.

The repeated cycles of erythrocytic infection, intracellular growth, rupture and re-infection give rise to clinical symptoms and pathology. Infected RBCs (iRBCs) and released parasitic material is thought to be involved in the induction of the host immune response. Malarial anaemia occurs as a result of RBC destruction and reduced RBC synthesis, which is partly due to the diversion of new erythroid progenitors towards myeloid/lymphoid development in the haematopoietic pathway (Belyaev et al., 2010). During development within iRBCs, *Plasmodium* parasites adapt their host cell in order to gain access to nutrients and export waste products. As a result, the surfaces of infected erythrocytes are extensively modified with parasite proteins. However, the presence of these molecules also makes iRBCs a target for innate and adaptive immune responses and clearance by the reticulo-endothelial system of the spleen.

1.2.2 Cyto-adherance

Patients experiencing severe malaria often have vascular occlusion by iRBCs, which may obstruct blood flow and cause pathology [reviewed by: (Rogerson et al., 2007, Mackintosh et al., 2004, Rasti et al., 2004, Van der Heyde et al., 2006, Dondorp et al., 2008, Maude et al., 2009a, Maude et al., 2009b)]. Whilst this has not been definitively proven as the cause of severe malaria, iRBCs are known to accumulate in the brains of patients dying from cerebral malaria (Montgomery et al., 2007), and cyto-adherant parasite phenotypes have a strong correlation with severe disease (Ockenhouse et al., 1991, Smith et al., 2000) [and reviewed by (Berendt et al., 1994, De Souza et al., 2009)].

Plasmodium iRBCs are known to cytoadhere to both endothelium and to uninfected RBCs; these phenomena are known as sequestration and rosetting, respectively. The formation of rosettes may conceal iRBCs from immune responses. Observations that *P. knowlesi* iRBCs no longer sequestered after host splenectomy (Barnwell et al., 1982, Barnwell et al., 1983) have led to the theory that sequestration of iRBCs in the deep vasculature prevents their clearance by the spleen. Several host receptors have been determined which mediate sequestration, including: cluster determinant 36 (CD36), vascular cell adhesion molecule-1 (VCAM-1), intercellular adhesion molecule-1 (ICAM1), E-selectin, platelet-endothelial cell adhesion molecule c1 (PECAM-1/CD31), P-selectin, chondroitin sulphate A (CSA) and hyaluronic acid (HA; (Barnwell et al., 1989, Ockenhouse et al., 1992, Berendt et al., 1989, Fried and Duffy, 1996, Rogerson et al., 1995, Treutiger et al., 1997, Beeson et al., 2000, Ockenhouse et al., 1989, Ho et al., 1998).

1.2.3 Models of malaria

Plasmodium parasites usually have a limited host range, and the five species infecting humans will only grow in higher primates or *in vitro* RBC culture for some species. This often makes dissection of key components for immunity to the human-infecting *Plasmodium* species unethical or inaccessible for researchers.

Fortunately, some of the features of human malaria are represented in rodent-infecting *Plasmodium* species, which may be used to model the human disease. Among these, *Plasmodium chabaudi* was isolated from *Thamnomys rutilans* thicket rats in the African

Congo (Landau, 1965, Landau and Boulard, 1978). Other rodent malarias include *P. berghei, P. yoelii* and *P. vinckei.*

Infection by different *Plasmodium* species can cause different pathologies - either acute (where the primary infection results in death or parasite clearance) or chronic (long-term) infections, which may recrudesce (forming secondary peaks of infection). The features of these rodent malarias are highly dependent upon the parasite strain used (McKenzie et al., 2008), in combination with the host genetic background, this is summarized in Table 2.

Table 2: Features of rodent malaria models.

Species, sub species, clone	RBC invasion pref.	Susceptibility to anaemia	Susceptibility to cerebral malaria	Organs of iRBC seq. (acc.)	References
P. berghei ANKA	RE	C57BL/6: L CD-1: L C57BL/6J: NL	C57BL/6: S CBA: S BALB/c: R	Lungs, Adipose tissue & Spleen (Brain)	(Franke-Fayard et al., 2005b, Engwerda et al., 2005, Asami et al., 1992, Maggio-Price et al., 1985, Franke-Fayard et al., 2006, Spaccapelo et al., 2010)
P. yoelii yoelii 17X	RE & E	BALB/c: NL	Most strains R	(Spleen)	(Asami et al., 1992, Landau and Boulard, 1978, Martin-Jaular et al., 2011)
P. yoelii yoelii 17XL	RE & E	BALB/c: L C57BL/6: L	Most strains S		(Weiss et al., 1989)
P. chabaudi chabaudi AS	E	A/J: L C57BL/6: NL	C57BL/6 IL-10 KO: S	Liver (Spleen & Brain)	(Gilks et al., 1990b, Mota et al., 2000, Yap and Stevenson, 1992, Sanni et al., 2004)
P. chabaudi adami DS	E	C3H: L C57BL/6: NL			(Villeval et al., 1990)
P. vinckei vinckei	E				(Silverman et al., 1987)

Adapted from: (Lamb et al., 2006). Preference was abbreviated to pref., reticulocytes and erythrocytes were abbreviated to RE and E, lethal and non-lethal were abbreviated to L and NL, resistant and susceptible were abbreviated to R and S, sequestration and accumuation were abbreviated to seq. and acc., respectively.

1.3 Protective immunity to malaria

1.3.1 Cells of the immune response

The immune system is divided into the innate and adaptive responses. Cells of the innate immune system comprise a host of leukocytes, which recognize and respond to pathogens non-specifically. These include mast cells, eosinophils and basophils, which produce histamine amongst their responses; Natural killer (NK) cells, which primarily target host cells that have down-regulated MHC class I, and the phagocytic neutrophils, macrophages, and dendritic cells. The latter two cell types are capable of presenting processed antigens to cells of the adaptive response, stimulating antigen-specific responses, and are thus known as antigen presenting cells (APCs).

The adaptive immune response is comprised of lymphocytes, all of which first develop in the bone marrow. T cells then migrate to the thymus, where they are exposed to the full host peptide repertoire and self-recognizing cells are deleted. Thymic development results in T cell subsets possessing the surface receptor CD8, also known as cytotoxic lymphocytes; and CD4, which are T helper (Th) cells.

Lymphocytes completing their development in the bone marrow are known as B cells. Once activated by a Th cell specific for the same antigen, B cells proliferate and differentiate into plasma cells, whose primary purpose is the production of antibody (Ab). Abs provide the basis for humoral immunity and consist of a variable, antigen specific region, and a constant region. The constant region provides different effector functions, according to which C-region gene segment is incorporated: IgA, IgD, IgE, IgG or IgM.

1.3.2 Immune response overview

Upon infection with a typical pathogen, the innate immune response is the first line of defence. Recognition of pathogen-associated molecular patterns (PAMPs) by pattern recognition receptors (PRRs) causes macrophages, dendritic cells and neutrophils to raise the alarm by secreting inflammatory cytokines and chemokines. This promotes migration of other leukocytes to the site of infection, and at the same time APCs migrate to lymphoid tissues to stimulate an antigen specific, adaptive response.

Antigen presenting cells process antigens into short peptides and present these in the context of MHC class II molecules to T helper cells [reviewed by (Jenkins et al., 2001)]. T helper cells recognizing their cognate antigen, in the presence of co-stimulatory molecules derived from the APC, are stimulated to divide and secrete different cytokines, depending upon the conditions under which they were activated [for example: (Hugues et al., 2004, Shakhar et al., 2005)]. The Th subsets are summarized in Figure 2.

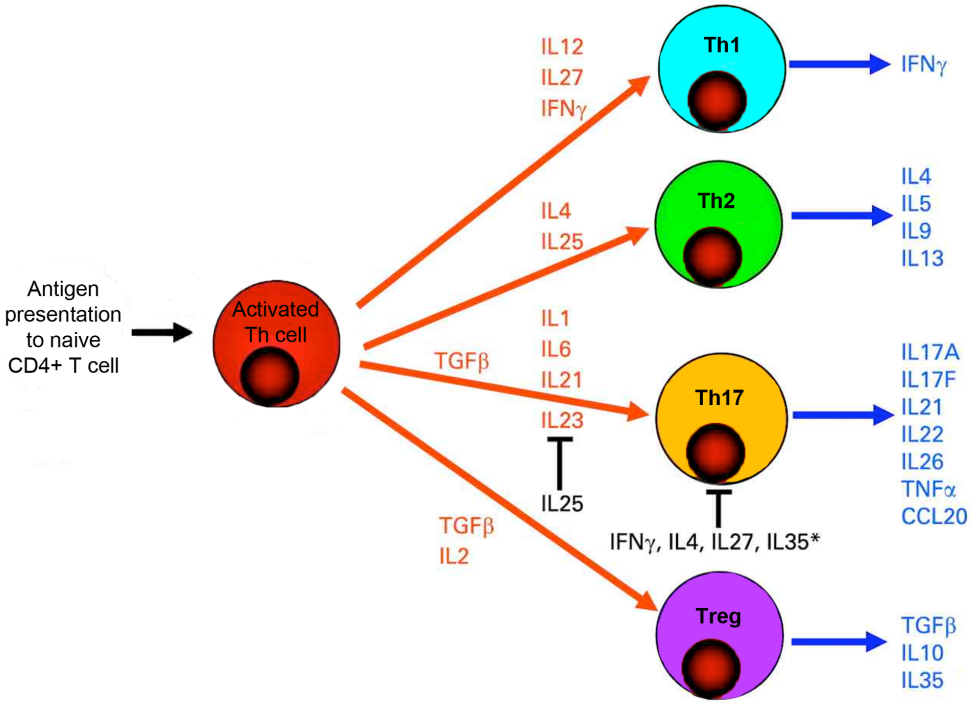


Figure 2: The major T helper cell subsets.

Adapted from (Brand, 2009). Naïve CD4+ T helper (Th) cells, once activated by presentation of their cognate antigen, differentiate into four major subsets: Th1, Th2, Th17 and T regulatory cells, which may develop within the thymus (Treg) or be induced in the periphery (iTreg). A fifth subset has been recently described, Th9 (Veldhoen et al., 2008), not included in this diagram. Different conditions dispose CD4 T cells to differentiate into the different subsets, indicated by red arrows and text. The cytokines produced by the different subsets are shown with blue arrows and text.

At first, B cells produce both IgM Abs and membrane bound IgD. The function of IgD is unknown, and animals lacking IgD appear to retain fully functional immune responses (Blattner and Tucker, 1984, Roes and Rajewsky, 1991). IgM Abs individually tend to bind antigen with low affinity, but high avidity; they can form pentamers which allows multi-point binding, and also efficiently activate the complement cascade, described below.

After B cell activation, the B cells and their cognate Th cells migrate into follicles within the lymphoid tissue where the B cells proliferate, forming germinal centres. Within the germinal centre further genetic refinement occurs via the processes of somatic hyper mutation and class switching. Somatic hyper mutation involves the accumulation of point mutations within the variable region, and enables more avid antigen recognition; whilst class switching involves physical recombination to change the constant region of the antibodies produced to either IgG, IgA and IgE Ab isotypes. These act in different tissues, IgG being the major isotype found in the blood and extra cellular fluid. IgG Abs may comprise one of four sub-classes, termed 1 - 4 in humans, and 1, 2a, 2b (BALB/c, or 2c in C57BL/6) and 3 in mice.

Antibody functions can be broadly divided into three types, examples of which are shown in Figure 3. All Ab isotypes can perform neutralization and opsonization functions to different degrees, whilst complement can be activated most efficiently by IgM, but also by IgG.

- Where pathogens are not recognized by pattern recognition receptors present on innate immune cells, antibodies can bind to the pathogen surface and signal for their phagocytosis, a process known as opsonization.
- Neutralization involves direct antigen binding, thus inhibiting pathogen invasion or the action of their products. This may be mediated by agglutination of Abs to each other, a process that can also aid opsonization.
- Complement is a cascade of plasma proteins that can directly kill certain pathogens, and can also promote opsonization. The complement cascade may be activated in a variety of ways, including the classical pathway by Ab: antigen complexes (Cooper, 1985).

The key advantage of the adaptive response over innate responses is the ability to generate memory of an infection. Memory T and B lymphocytes can divide and respond more quickly and effectively to a subsequent infection, whilst the residual presence of long-lived plasma cells in bone marrow provides a threshold level of serum-antibody that should prevent establishment of a subsequent infection.

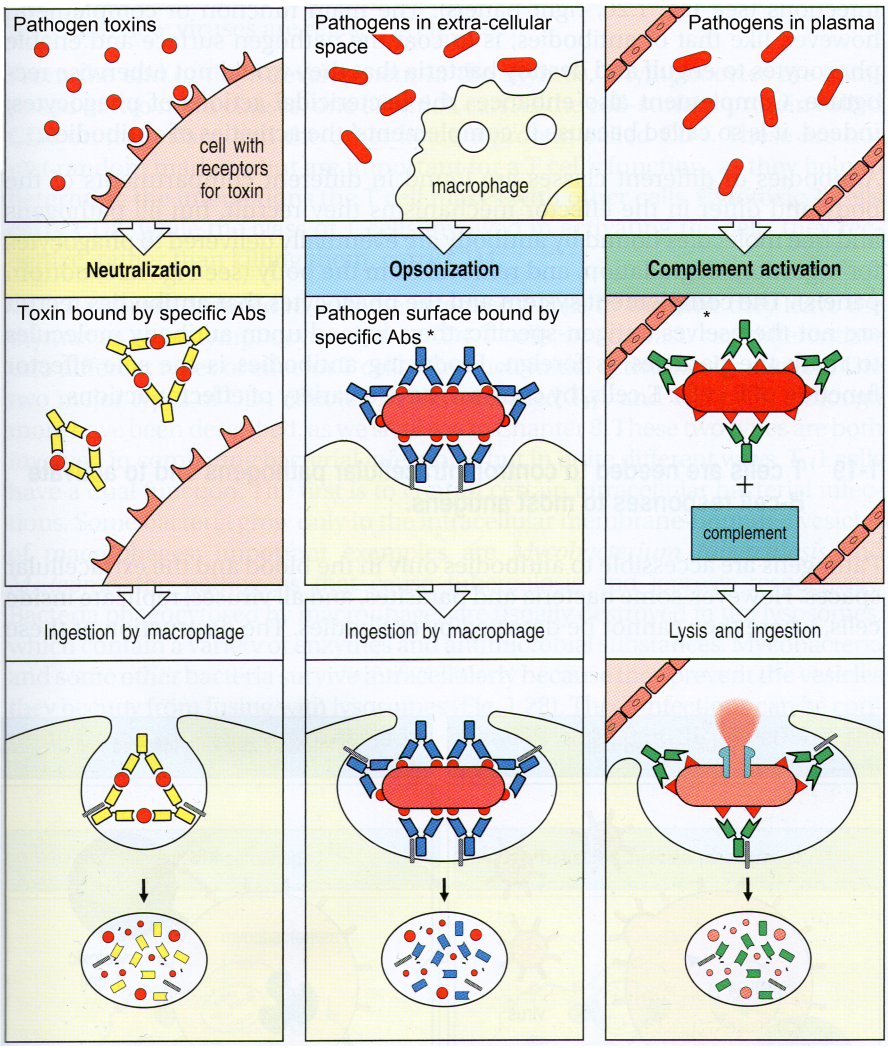


Figure 3: The three major functions of antibodies.

Adapted from (Janeway, 2007). These examples show three of the possible antibody functions during bacterial infection, although the same responses also occur during *Plasmodium* infection. The left panels show neutralization of a bacterial toxin, which may then be ingested by macrophages. This could equally apply to *Plasmodium* toxins, such as glycosylphosphatidylinisotol. The centre panels show opsonization of whole bacteria, followed by phagocytosis, this process may occur for merozoites or iRBC during the erythrocytic stages of *Plasmodium* infection. Finally, the right panels show the binding of complement to the bacterial cell, followed by direct complement-mediated lysis. This mechanism acts on *Plasmodium* gametocytes (Healer et al., 1997), however *in vivo*, iRBC appear to be protected from complement mediated lysis (Kawamoto et al., 1997, Wiesner et al., 1997).

1.3.3 Defining immune responses to malaria

Selective experimental removal of cell populations (either by depletion or the use of knock-out animals) has allowed investigation into which cell types are essential for resolving *Plasmodium* infection and providing immune memory. Many of these studies have used intra-peritioneal infection of mice with *P. chabaudi*, a model for the blood stages of malaria.

In this model, C57BL/6 and BALB/c mice experience a peak of parasitaemia around day 7 when injected with an infectious dose of 10^5 iRBCs. At the peak, 20-40% of RBCs may be infected; this is quickly controlled, and develops into a chronic infection where parasites may remain at sub-patent levels for several months (Wargo et al., 2006). Recrudescences occur on day 25 and around day 30, each comprising iRBCs with different surface phenotypes than those observed previously in the infection (McLean et al., 1982). A typical course of parasitaemia is illustrated in Figure 4, along with immune characteristics important for resolution of infection.

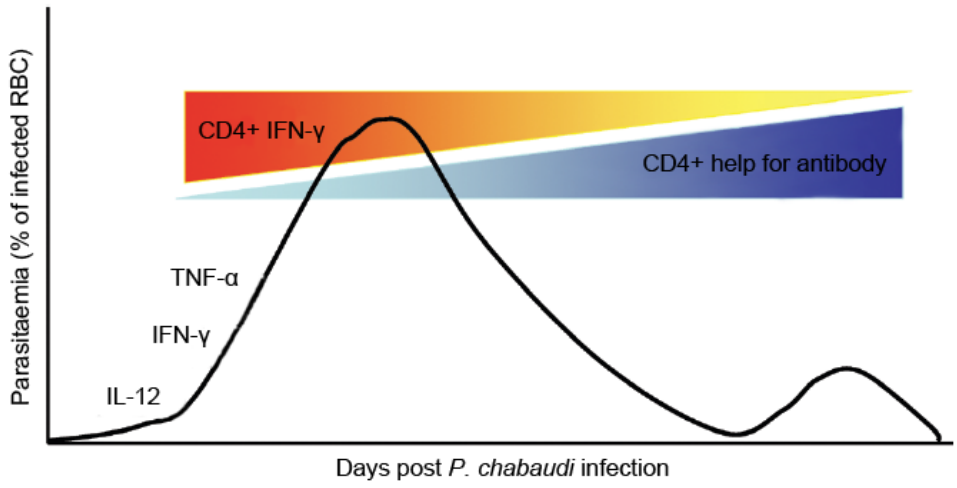


Figure 4: Characteristics of *P. chabaudi* infection in mice.

Adapted from (Langhorne et al., 2004). The percentage of RBC infected is shown in black. During the first few days of infection, interleukin 12 is produced by macrophages and dendritic cells, thereby inducing development of a Th1 response, characterized by high levels of interferon-γ (IFN-γ) and tumour necrosis factor-α (TNF-α). This inflammatory environment aids control of peak parasitaemia around day 10 of infection. To prevent immune-mediated pathology and provide help for Ab production, the CD4 positive population then switches to become predominantly Th2 type, allowing the production of the high Ab titres required for parasite clearance.

During the initial stages of infection, cells of the innate response act in various ways to

inhibit parasite growth, including the release of parasitocidal mediators and receptor-mediated phagocytosis by macrophages and monocytes (Stevenson and Riley, 2004). Other innate cells such as NK, natural killer T- (NKT), and gamma delta T- (γδT) cells may also be important for early control of parasitaemia (Mohan et al., 1997, Seixas et al., 2002, Artavanis-Tsakonas and Riley, 2002, Roetynck et al., 2006).

Resolution of *P. chabaudi* infection, however, requires both T and B lymphocytes: in their absence, mice were unable to control the peak of parasitaemia and succumbed to infection. Equally, when CD4+ cells were adoptively transferred into these animals, they were able to control the acute phase of infection but not to eliminate the parasites (Meding and Langhorne, 1991).

Control of the peak of infection is characterized by inflammation, predominantly high levels of interferon gamma (IFN-γ), and the signalling cascade that it induces (Meding et al., 1990, Stevenson et al., 1990, van der Heyde et al., 1997, Favre et al., 1997, Tsuji et al., 1995). This is typical of a Th1 response, which is known to activate macrophages. Indeed, it has recently been shown that inflammatory monocytes, which differentiate into macrophages, aid parasite clearance around day 10 of infection (Sponaas et al., 2009). After the acute phase of parasitaemia is resolved, CD4 responses convert towards a Th2 phenotype, producing IL-4, IL-5, IL-6, and IL-10, and stimulating B cells to produce antibody (Langhorne, 1989). Importantly, in the absence of B cells, mice were able to control the peak of parasitaemia to sub-patent levels but unable to eliminate parasites (Grun and Weidanz, 1983).

It is likely that all of the three mechanisms described in Figure 3 are involved in parasite removal. In *P. chabaudi*, the absence of complement did not affect the ability of mice to resolve primary infection, but increased susceptibility to re-infection (Taylor et al., 2001). Evidence from *P. falciparum* implies that neutralizing Abs inhibit RBC invasion (Blackman et al., 1990), and that opsonization aids iRBC clearance from both blood and placenta, although it is not required for complete parasite elimination (Bull et al., 1998, Keen et al., 2007, Rotman et al., 1998). Abs generated to antigens displayed on the iRBC surface enhance phagocytosis and destruction of iRBC *in vitro* (Mota et al., 1998), and the presence of opsonising Abs is associated with protection (Groux and Gysin, 1990). Importantly, passive transfer of Abs from clinically immune adults has been successfully used to treat patients with severe malaria (Cohen et al., 1961, Sabchareon et al., 1991), demonstrating the crucial role of Abs in immunity to severe disease.

1.3.4 Generation of immune memory to *Plasmodium*

The acquisition of immunity to malaria in humans is complicated by many factors, including infection with different *Plasmodium* species and strains, host genetic variation and transmission intensity. For example, both in Kenya and the Gambia, the risks of developing cerebral malaria or severe malaria anemia in childhood were reported to be lowest among populations experiencing the highest transmission intensities, and vice versa (Mbogo et al., 1993, Mbogo et al., 1995, McElroy et al., 1997, Snow et al., 1997).

Nonetheless, malaria infection follows a common pattern in most endemic areas of sub-Saharan Africa: Infants are initially protected from parasitaemia, and severe disease, probably due to the presence of maternal IgG. Around 3-4 months old, children begin to experience clinical episodes and are extremely vulnerable to severe disease and mortality. Cerebral malaria (CM) commonly develops in children between 2 and 4 years of age, the risk of this complication usually subsides after age 5. After puberty, severe disease is rare, although clinical episodes may occur several times per year. This situation is maintained throughout adult life in endemic areas, and many people appear to develop immunity to clinical disease, whilst maintaining asymptomatic parasitaemia (Doolan et al., 2009).

There is an important exception to the above: pregnancy associated malaria [PAM, (Rogerson et al., 2007, Hviid et al., 2010)]. Women otherwise clinically immune to malaria may experience severe disease, predominantly during their first pregnancy, after which immunity to the condition begins to develop. This is caused in part by the accumulation of iRBCs in the placenta, reducing birth weight or leading to miscarriage.

For reasons awaiting discovery, individuals living in malaria-endemic areas appear never to develop protection from infection, otherwise known as sterilizing immunity (Sergent and Parrot, 1935, Obi et al., 2011). Therefore, naturally acquired immunity to pre-erythrocytic stages has been considered ineffective (Hoffman et al., 1987). In contrast, deliberate exposure to pre-erythrocytic antigens by vaccination is capable of inducing protective responses (Mueller et al., 2005, Hoffman et al., 2002, Alonso, 2006).

1.4 Factors influencing the outcome of malaria

The generation of protective immunity and the severity of disease during *Plasmodium* infection may be affected by several factors, which are considered below.

Firstly, epidemiological data suggest that in endemic areas, almost a quarter of the risk of developing severe malaria is due to host genetic factors (Mackinnon et al., 2005). For example, many protective traits affect the ability of parasites to invade and replicate within RBCs, and as such have been preserved in populations continually exposed to malaria infection, whilst these factors have been removed by selection in non-malaria exposed populations. Examples include sickle cell trait, thalassaemia, glucose-6-phosphate deficiency and ovalocytosis, which provide protection for heterozygous individuals (Allen et al., 1997, Gilles et al., 1967, Miller, 1999, Ruwende et al., 1995). Also the Duffy negative trait common in African peoples prevents RBC invasion by *P. vivax* (Miller et al., 1976), resulting in Africans comprising less than 4% of the global population at risk of *P. vivax* infection (Guerra et al., 2010). Considerable host genetic variability also lies in the composition of each individual's immune response. For example, whilst inflammatory Th1 responses are required to control parasitaemia and for generation of protective immunity (Sam et al., 1999, Su et al., 2002), over-production of inflammatory cytokines such as TNFα and IFNγ cause immuno-pathology (Kwiatkowski et al., 1990, Lyke et al., 2004). The regulatory cytokine IL10 reduces inflammation, to the extent that *P. chabaudi* infection in mice lacking IL10 is frequently lethal (Li et al., 1999). Therefore polymorphisms in the genes or regulation of IL10, TNFα and IFNγ (Juliger et al., 2003, Carpenter and al., 2007, Bayley et al., 2004, Henri and al., 2002) could play a critical role in the ability to clear infection, and the extent of immuno-pathology experienced (Li et al., 2003, Omer et al., 2003, Kurtzhals, 1999).

Secondly, people who leave malaria-endemic areas appear to lose their previously established immunity, such that upon moving back to the endemic area, they are equally susceptible to infection as naïve individuals (Jennings et al., 2006). This has led to conjectures that immune memory may be somehow defective - a hypothesis investigated in our laboratory using antigen specific responses to Merozoite Surface Protein 1 (MSP1) in the model *P. chabaudi*, outlined below:

- The generation of humoral responses appeared intact, as MSP1 specific memory B cells and IgG secreting long-lived plasma cells were detectable for more than eight months following primary infection. Upon homologous challenge, functional memory B-cell and long-lived plasma cell responses were induced (Ndungu et al., 2009). These memory B cells could be detected in peripheral blood throughout infection, but plasma cells were only transiently detected in the blood during acute infection, possibly reflecting their migration to bone marrow (Nduati and Ng, 2010). These data support the finding that *P. falciparum* specific memory B cells can persist in adults in the absence of exposure for over eight years (Migot et al., 1993), and suggest that deficiencies in the generation and maintenance of the humoral response do not contribute to prolonged susceptibility to malaria.

- CD4+ T cells specific for malarial antigen MSP1 have been shown to protect immuno-deficient $RAG2^{-/-}$ mice from an otherwise lethal *P. chabaudi* infection, providing that a threshold of anti-parasite Abs was also present (Stephens et al., 2005). After experiencing a primary *P. chabaudi* infection, these CD4+ T cells could be visualized in three subsets: central memory, and early and late effector memory cells, based on the relative expression of the following surface markers: CD44, IL-7Rα, CD62L and CD27, and the cytokines IFNγ and IL10 (Stephens and Langhorne, 2010). Adoptive transfer of each of these sorted subsets, isolated from a chronically infected mouse, was able to protect $RAG2^{-/-}$ mice both from excessive parasitaemia and pathology, better than the same cells isolated from drug-cured mice (Stephens and Langhorne, 2010). These data suggest that chronic antigen stimulation aids protective memory CD4+ responses.

In the field, humoral immunity generated to MSP1 after infection correlates with substantial protection against homologous challenge (Weir and Cockerham, 1984, Riley et al., 1992, Egan et al., 1996, Su and Wellems, 1996, Branch et al., 1998, Anderson et al., 1999, Dodoo et al., 1999, Kitua et al., 1999, Conway et al., 2000, Metzger et al., 2003, Polley et al., 2004, Perrant et al., 2005, Nebie et al., 2008, Osier et al., 2010, Stanisic et al., 2009). Whilst it is possible that responses to MSP1 may not be representative of the total immune response to other antigens, these data suggest that established immune memory to *Plasmodium* can persist. This is supported by evidence from Madagascar, where malaria was eradicated for 30 years, before re-emergence.

Adults who were immune before eradication remained more resistant to clinical disease than naïve individuals (Deloron and Chougnet, 1992).

Data acquired from field studies where *P. falciparum* infection was dominant (Marsh and Kinyanjui, 2006) indicate that there are three phases of susceptibility to malaria, which are shown in Figure 5.

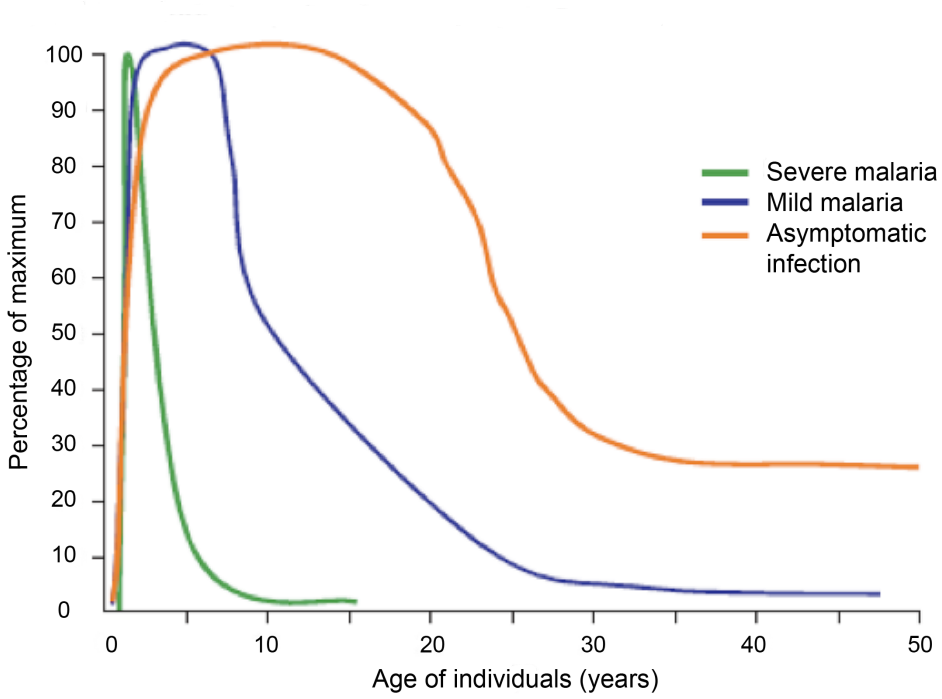


Figure 5: The different types of malaria experienced by endemic populations.

Adapted from (Marsh and Kinyanjui, 2006). The percentage of individuals in the endemic area of Kilifi, Kenya, experiencing severe malaria (green), 'mild' uncomplicated malaria (blue), or asymptomatic *Plasmodium* infection (orange), are shown, along with the ages at which these syndromes manifest.

Immunity to severe disease and morbidity and may be acquired fairly quickly in life. Gupta and colleagues used mathematical modelling to suggest that this phase of immunity (non-cerebral severe malaria defined as "*Plasmodium falciparum* infection requiring hospital admission") may be acquired after only one or two infections (Gupta et al., 1999). Immunity to clinical symptoms takes longer to develop and is often negatively associated with high parasite load during infection (Marsh and Snow, 1997). The acquisition of immunity to clinical malaria nonetheless allows chronic *Plasmodium* infection to occur, resulting in asymptomatic parasitaemia (Sergent and Parrot, 1935, Koch, 1900a, Koch, 1900b, Koch, 1900c, Koch, 1900d).

There are two plausible reasons for this three-phase pattern of *Plasmodium* infection:

i) Qualitatively different immune responses are required at each stage of malaria.
ii) Certain parasite phenotypes are disposed towards causing infection with different characteristics. In this case immunity would need to be generated towards different parasite strains or variants, in order to prevent their associated pathologies.

There is evidence to support both theories, and they may not be mutually exclusive. Whilst the content of qualitatively different immune responses is not clear, different states of immuno-modulation appear to associate with severe and uncomplicated disease. Severe malarial disease syndromes are often associated with high levels of inflammation, giving rise to immuno-pathology [reviewed by (Mackintosh et al., 2004, Combes et al., 2010, Moxon et al., 2010)]. Endothelial cells respond to inflammation by up-regulating expression of receptors such as ICAM-1 and VCAM-1, to which iRBCs can cyto-adhere. This response is more pronounced in endothelium from patients with severe disease than uncomplicated malaria (Wassmer et al., 2011). A further mechanism which may contribute to severe disease has been recently been put forward, whereby activated endothelial cells produce long strings of von Willebrand Factor, to which platelets and iRBCs can adhere, thus enhancing the capacity of iRBCs to sequester (Bridges et al., 2010).

The inflammation state affects other features, such as the activation status of platelets and intravascular accumulation of leukocytes, both of which are complications during cerebral malaria [for example (Grau et al., 2003, Taylor et al., 2004, Medana and Turner, 2006, Miu et al., 2008)]. In addition, different forms of severe malaria have been linked to particular cytokine profiles (Prakash et al., 2006); for example IL10 can be associated with either severe anaemia, at low levels (Kurtzhals et al., 1998), or respiratory distress, at high levels (Awandare et al., 2006). Factors such as prior exposure to *Plasmodium* infection, nutritional status and co-infection context may influence an individual's inflammation status and cytokine profile during different malaria episodes.

An alternate explanation for the tendency towards severe or uncomplicated disease has been put forward regarding parasite retention by the splenic architecture (Safeukui et al., 2008). In this argument iRBCs not retained within the spleen would have the potential to sequester in different organs, causing severe disease; whilst if the majority

of parasite biomass were retained in the spleen, fewer iRBCs would have the capacity to sequester. The latter scenario would result in splenomegaly, and probably contribute to malarial anaemia, yet would allow the infection to progress in an uncomplicated fashion. The factors influencing such splenic retention are unclear, but it is likely that possession of immunity to a previous infection would enhance the ability of the spleen to behave in this way. Buffet and colleagues have proposed that the splenic retention of iRBCs is more stringent in infants, allowing their spleens to retain more iRBCs during a primary infection, than the spleens of naïve adults (Buffet et al., 2011). This may explain observations in S. E. Asia that naïve adults are more susceptible to severe disease than naïve young children [for example (Baird et al., 2003)]. However, immunity to *P. falciparum* may not be acquired in the same way as immunity to *P. vivax* and other species prevalent in different regions. Immune memory may also not be acquired in the same way in areas of high and low or seasonal transmission, making parallels difficult to draw between different populations. Another factor is that children may have a different capacity to generate immunity than adults. For example, children develop short lived Ab responses following infection and fail to boost Ab titres upon re-infection (Kinyanjui et al., 2007). It has been suggested that the onset of puberty aids generation of widespread immunity to malaria (Kurtis et al., 2001).

Whilst different immune responses could contribute to the three phases of malaria shown in Figure 5, they do not explain observations of strain specific immunity. Primary malaria infection results in protective immunity to homologous but not heterologous challenge in monkey and rodent models (Cadigan Jr and Chaicumpa, 1969, Jones et al., 2000, Jarra and Brown, 1985). This has also been observed in experimental human infections (Pombo et al., 2002, Krause et al., 2007), and when malaria infection was used as a treatment for patients suffering from neurosyphilis (Yorke and Macfie, 1924, Ciuca et al., 1934, James and Ciuca, 1938, Jeffery, 1966).

From these data, the dominant hypothesis has emerged that individuals become immune to clinical symptoms only after being exposed to the majority of locally circulating *Plasmodium* strains (Garnham, 1966, Wilson and Phillips, 1976, Day and Marsh, 1991, Bruce et al., 2000). This could be due either to polymorphism in conserved antigens (Conway et al., 2000, Healer et al., 2004, Coley et al., 2006, Hodder and al., 2001, Kennedy et al., 2002, Miura et al., 2007, Tanabe et al., 1987, Polley and Conway, 2001), or the differential expression of antigens, including members of diverse multi-gene families.

1.5 Protective immunity against parasite variants

Studies undertaken in Kenya have shown that sera taken from children and adults have differential abilities to recognize and agglutinate iRBCs isolated from different *Plasmodium* infections. Typically, a child's serum could only agglutinate parasites isolated from their own infection, whilst sera from immune adults could agglutinate a range of parasite isolates (Bull et al. 1999, 2000). In addition, parasites isolated from severe malaria cases were more likely to be recognized by sera from semi-immune children and adults, having a higher agglutination frequency, depicted in Figure 6, (Bull et al., 2000). These observations support the hypothesis that immunity to different parasite strains is the reason for increased protection from severe disease or clinical symptoms.

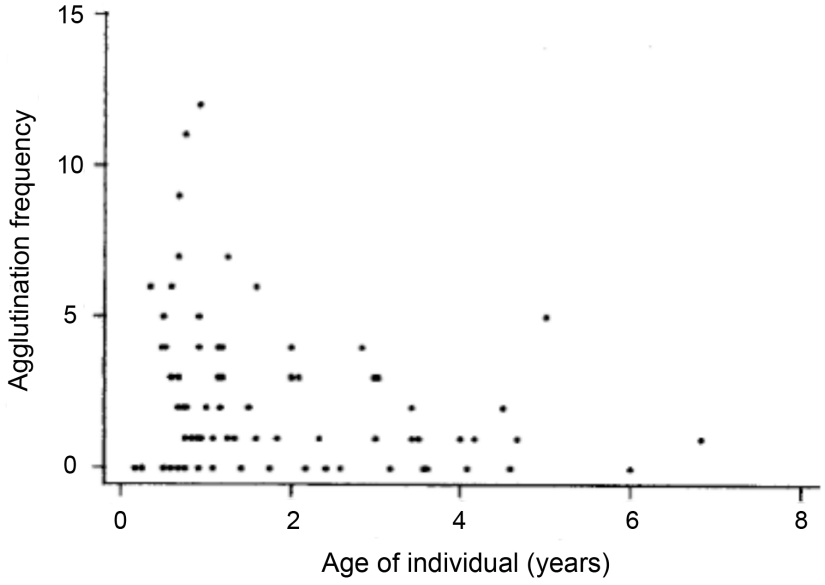


Figure 6: Frequency of iRBC agglutination by host immune serum, according to the age at which parasites were isolated from infected children.

Adapted from (Bull et al., 2000). *P. falciparum* isolates were taken from 70 children, in Kilifi, Kenya. Each isolate is plotted as a single point according to the age of the child from whom it was taken, shown on the X-axis. The frequency of agglutination of these parasites by the sera of all 70 children is shown in the Y-axis.

Because children develop immunity to severe disease fairly rapidly, as discussed above, the parasite variants encountered during these early infections must be associated with severe disease. Therefore, the acquisition of immunity to different parasite antigens and their variants must be important.

There is no evidence that particular alleles of polymorphic antigens, such as apical membrane antigen 1 (AMA1) and MSP1, are associated with severe disease. Although immunization with these antigens is able to induce protection against re-infection in model systems (Daly and Long, 1993, Ling et al., 1997, Crewther et al., 1996, Hirunpetcharat et al., 1997, Tian et al., 1997, Anders et al., 1998, Burns et al., 2003, Burns et al., 2004, Alaro et al., 2010), recent phase II vaccine trials have not shown particular efficacy (Ogutu et al., 2009, Sagara et al., 2009). This suggests that acquisition of immunity to conserved antigens, whilst important for control of parasitaemia, is not the crucial factor determining the characteristics of malaria infection - whether severe, uncomplicated or asymptomatic.

It is important to note that very few *Plasmodium* antigens have been studied so far, approximately 1% from a predicted proteome of over 5000 proteins (Gardner et al., 2002, Langhorne et al., 2008). Scant success has been achieved to date using conserved antigens in vaccine trials, highlighted by the number of discontinued vaccines in Table 3. Clearly, there is an urgent need to identify new potential vaccine targets: antigens that are associated with parasites causing severe disease and to which protective immune responses are directed.

Table 3: A summary of current vaccine studies.

Antigen (s)	Pre-clinical	Clinical	Discontinued
Pre-erythrocytic stages			
Circumsporozoite protein (CSP)	3	4	8
Liver stage antigen 1	0	0	2
Thrombospondin related anonymous protein (TRAP)	0	0	4
Liver stage antigen 3	0	0	3
CSP & another antigen	1	1	5
Whole sporozoite antigens	1	1	0
Erythrocytic stages			
Apical membrane antigen 1	1	3	4
Merozoite surface protein 1	1	2	3
Merozoite surface protein 2	0	0	1
Glutamate rich protein	0	0	1
Merozoite surface protein 3	0	2	0
Erythrocyte binding antigen 175	0	1	0
VAR-2-CSA	2	0	0
Combinations of these antigens	1	0	3
Sexual stages			
P. falciparum surface protein 25 (*Pf*s25)	1	0	1
P. vivax surface protein 25	0	0	1
*Pf*s25 & another antigen	1	0	0
Whole gametocyte antigens	0	2	0
Multiple stages	1	1	1

Adapted from from WHO data 2009: www.who.int/malaria/world_malaria_report_2009/en/index.html

Support for the idea that particular parasite variants are associated with severe disease comes from typical parasitaemia courses observed in human and model infections, whereby after resolution of the acute phase, primary infections are not resolved for some time. Instead, several recrudescences occur, an example of which is displayed in Figure 7. In this experimental human infection, there were repeated peaks of infection, and although the intensity of parasitaemia diminished over time, the infection had still not been cleared by day 260, whereupon drug cure was administered [described by (Miller et al., 1994)].

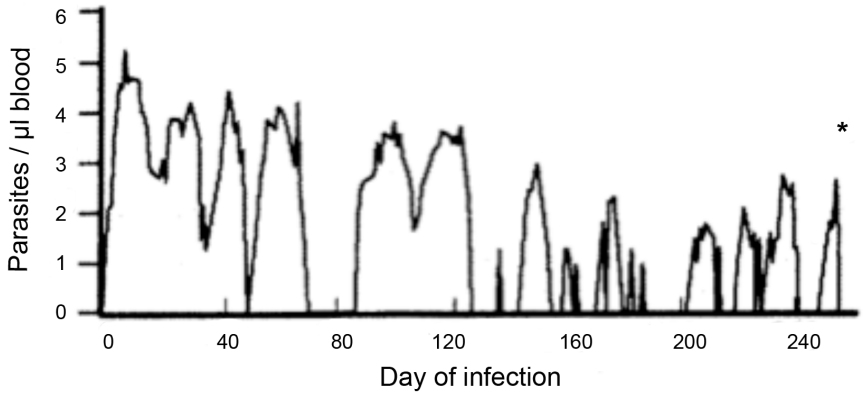


Figure 7: Experimental *P. falciparum* infection of a naïve human volunteer.

Adapted from data described in (Miller et al., 1994). Peripheral parasitaemia, defined as the number of iRBC per µl blood x 10^6, was monitored during the course of infection, until chloroquine drug-cure was administered on day 260 of infection, denoted by *.

The crucial factor in this example is that the experimental infection was known to be a single parasite strain, inoculated via mosquito bite, yet repeated immune responses were evaded for a considerable length of time (Miller et al., 1994). This could not be due to induction of expression of different alleles of a conserved protein, because the parasite genome could not evolve so rapidly. Instead, this points to the differential expression of variants encoded within the genome, each of which could be novel to the ongoing immune response, allowing continued parasite growth.

1.6 Antigenic variation

The capacity to change surface antigens in the face of dynamic host immune responses, is an advantage for pathogens, and has been termed antigenic variation [reviewed by (Deitsch et al., 2009, Turner, 2002)]. This phenomenon does not include permanent genetic change and probably occurs by immune selection of spontaneously changed surface antigens (Roberts et al., 1992, Brannan et al., 1994).

Two types of antigenic variation have been described. Phase variation refers to the ability to switch antigen expression on or off. This was first described for flagellar expression by *Salmonella* (Zeig et al., 1977), and is in fact a mechanism common to all pathogens. Antigenic variation in the true sense refers to the ability to express alternative forms of an antigen. This can occur within a gene by alternative splicing events, or by expression of different members of multi-gene families. Protozoan parasite genomes contain diverse multi-gene families, examples are shown in Table 4.

Table 4: Comparison of variant antigens encoded by protozoan multi-gene families.

Parasite	***Trypanosoma brucei***	***Giardia lamblia***	***Babesia bovis***
Lifestyle	Initially blood and haemolymph dwelling, followed by CNS* invasion.	Intestinal dwelling trophozoites	Sporozoites directly infect RBC
Host	Human	Human	Cattle (human zoonosis)
Trans-mission	Tsetse fly	Environmentally resistant infectious cysts	Tick
Disease	African sleeping sickness (headache, fever, swollen lymph nodes and neurological symptoms)	Abdominal cramps, diarrhoea and nausea	Fever, malaise, and hemolytic anemia
Gene family	Variant surface glycoprotein (VSG)	Variant surface protein (VSP)	Variant erythrocyte surface antigen 1 (VESA1)
Copy no.	≥ 1250-1500	> 190	~130-160
Regulation of gene expression	Partial genes recombine into a single active expression site, whose activity is epigenetically regulated	Several *vsp* genes are transcribed but 'silent' transcripts are degraded via the RNAi pathway.	The two subunits are encoded by the *ves1a* and *b* genes, whose expression is regulated like the *vsg* genes.
Other	Hydrodynamic flow directs Ig-VSG complexes to the cell posterior where they are endocytosed, allowing growth in the face of low-level humoral responses.	Disruption of the RNAi pathway allows all VSP to be expressed simultaneously, permitting generation of effective immune memory.	VESA1 proteins are located in iRBC membrane knobs and mediate both cytoadhesion to endothelium and antigenic variation
References	(Bernards et al., 1981, Cross, 1996, Engstler et al., 2007) *Central Nervous System	(Prucca et al., 2008, Rivero et al., 2010)	(Allred et al., 2000, Al-Khedery and Allred, 2006, Xiao et al., 2010)

Estimates of copy number downloaded from http://www.vardb.org/vardb/families.html, March 2011.

The major benefit of antigenic variation is that pathogens can evade immune recognition, which allows long-term (chronic) infection. It is important to note that antigenic variation may not be the primary function of such proteins, and in fact, varying their expression in this manner may act to preserve their real function.

For *Plasmodium* parasites, chronic infections are necessary to increase their chances of transmission into another mosquito host and continue the life-cycle. Other benefits include the ability to re-infect the same hosts such that the parasite-host range is not decreased, and even to super-infect hosts already carrying a different *Plasmodium* infection. This is certainly true for *P. falciparum*, where individuals can be infected with 7 genetically distinct parasite populations (Montgomery et al., 2007).

Variant antigens have been shown to be important targets of protective immunity (Marsh and Howard, 1986, Marsh et al., 1989), and crucially, certain variants are associated with different disease states (Bull et al., 2000a, Bull et al., 2008, Nielsen et al., 2002, Kirchgatter and Del Portillo, 2002, Kyriacou et al., 2006a, Kaestli et al., 2006, Rottmann et al., 2006, Normark et al., 2007). In the absence of pre-established immunity, parasites express these variants first so these are the proteins children initially make Abs against (Cham et al., 2009). Over the course of repeated infection, children acquire a repertoire of Abs against the variants that they have encountered, which correlates with the age that immunity to severe disease is acquired (Cham et al., 2009). This is shown in Figure 8, using data acquired from a village experiencing medium transmission intensity (Cham et al., 2009).

The cumulative acquisition of a broad range of Abs to an antigen expressed at the surface of iRBCs, *P. falciparum* Erythrocyte Membrane Protein 1 (*Pf*EMP1), is independent of malaria endemicity, although the pace of acquisition reflects the transmission intensity, with the acquisition of Abs recognizing more variants occurring at a younger age in areas of higher transmission (Cham et al., 2009). This would also correlate with the phenomenon observed in the field, that it takes much longer to develop immunity to clinical symptoms than to severe disease; due to the relatively restricted set of variants associated with the severe phenotype, compared with the much broader repertoire of less pathogenic variants which the parasite is obliged to express to evade previously established immunity (Marsh and Howard, 1986).

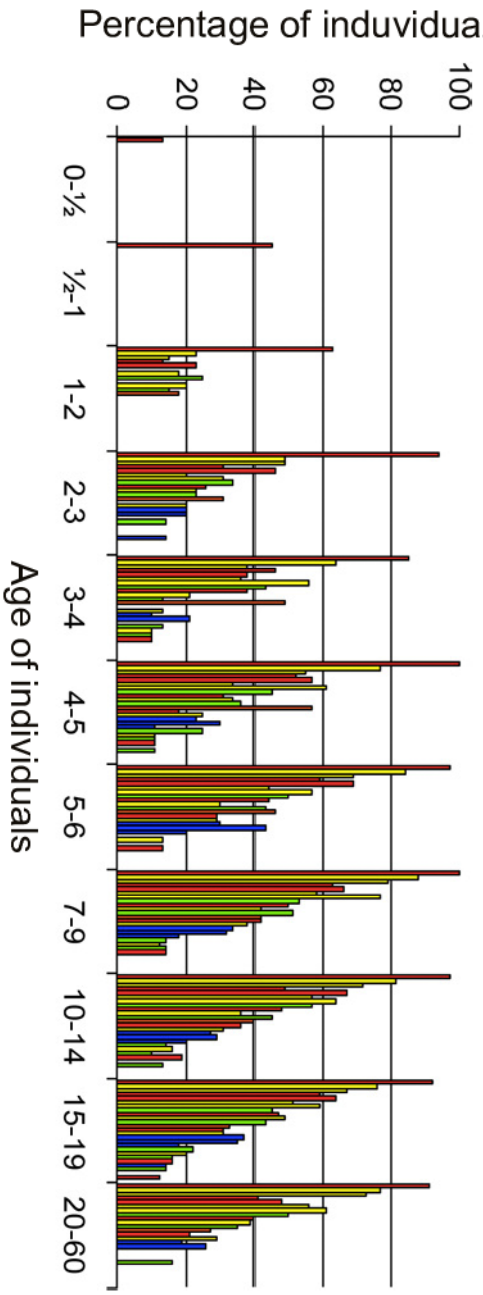


Figure 8: The percentage of individuals whose anti-sera recognized individual recombinant domains from the different *Pf*EMP1 sub-groups.

Adapted from (Cham et al., 2009). Serum samples from 1342 individuals were probed for recognition of 25 different recombinant *Pf*EMP1 DBL domains.

The study population was based in Tanga, Tanzania, where different villages experience variable intensities of malaria transmission. The data shown here were from a population experiencing high transmission frequency, and were classified according to the age of indviduals.

The percentage of individuals possessing IgG Abs recognizing each DBL domain is shown, for domains where >10% of individuals had a measurable response.

Each of the 25 *Pf*EMP1 DBL domains is shown as an individual column, coloured according to the native *Pf*EMP1 group: A, red; B/A, yellow; B, C, and B/C, green; VAR1, blue.

1.7 Variant surface antigens in *Plasmodium*

Several multi-gene families have been described in *Plasmodium*, some of which have been implicated in antigenic variation. These are introduced below.

The SICAvar proteins were the first family of variant surface antigens (VSAs) discovered in *Plasmodium* (Brown and Brown, 1965), using an assay to measure agglutination of infected red blood cells by antiserum from *P. knowlesi* infected rhesus macaques, the Schizont-Infected Cell Agglutination test [SICA, (Eaton, 1938)]. During *P. knowlesi* infection, each emergent parasite population was shown to have surface antigens that were distinct from those previously expressed. The genes responsible were later shown to be encoded by the *SICAvar* gene family (Al-Khedery et al., 1999). In fact the *SICAvar* genes comprise the largest gene family encoding variant antigens in *P. knowlesi*, with an estimated 107 members in the H strain (Pain et al., 2008). The predicted proteins range in size from 53 to 247 kDa (Pain et al., 2008).

Several variant antigen-encoding gene families have been identified in *P. falciparum*, in contrast to other *Plasmodium* species. The best-described family are the *var* genes, which encode the *Pf*EMP1 antigen. Members of the *Pf*EMP1 family are large proteins (200-350kDa) found in electron-dense knobs on the iRBC surface. There are approximately 60 *var* genes per parasite (Gardner et al., 2002), and the genes have a two-exon structure, the first exon encoding the extracellular domain (Su et al., 1995). All *Pf*EMP1 variants have several domains in the encoded amino acid sequence: an N-terminal segment; duffy binding like domains (DBL) with subtypes: α, β, γ, δ, ε and x; cysteine rich inter-domains (CIDR) with subtypes: α, β, γ; a trans-membrane (TM) domain, a C2 domain and a conserved intracellular acidic terminal segment (Kraemer and Smith, 2003), shown in Figure 9.

Between different *P. falciparum* isolates, the *var* genes are highly variable but while they may be re-arranged, the domains themselves are conserved (albeit at low sequence identity). Kraemer and colleagues described 31 different *Pf*EMP1 architectures, 7 of which were shared between all three sequenced isolates: 3D7 (Africa), HB3 (Central America) and IT4 (S. E. Asia) (Kraemer et al., 2007).

Figure 9: The arrangement of domains in an example *Pf*EMP1 molecule.

Adapted from (Kyes et al., 2007). A typical B or C-type *var* gene is shown with the encoded *Pf*EMP1 domains arranged from the N-terminal segment (NTS), followed by CIDR1α, DBL2β, C2, DBL3γ, DBL2β, CIDR2β, and TM domains, with a C-terminal acidic terminal segment (ATS). Within the 5' flanking DNA sequence, the upstream (Ups) region by which *var* genes are classified is denoted by *.

The chromosomal position, domain structure and upstream (Ups) DNA sequences of *Pf*EMP1 molecules define 4 sub-groups (Kraemer and Smith, 2003, Lavstsen et al., 2003, Gardner et al., 2002). Group A is the most complex, located sub-telomerically and transcribed towards the telomere, this group also incorporates *var* genes formerly named group D (Kraemer et al., 2007). A-type *Pf*EMP1 molecules contain an N-terminal DBL1α1 domain rather than the DBL1α domain common in other *Pf*EMP1sub-types. The domain arrangement of a typical B or C type *Pf*EMP1 protein is depicted in Figure 9. Group B *var* genes are transcribed away from the telomere and may be located subtelomerically or centrally, whilst group C are only found in central-chromosomal locations. Group E is located and transcribed similarly to group A, but includes only the most highly conserved *var*, including the pregnancy associated variant *var2csa.*

The majority of *var* genes are located in sub-telomeric clusters, where frequent recombination events take place, generating further diversity (Mu et al., 2005, Mu et al., 2007). Both reciprocal recombination and gene conversion events have been experimentally detected in the *var* repertoire (Su et al., 1999, Taylor et al., 2000b, Freitas-Junior et al., 2000). Although each parasite contains approximately 60 *var* genes, there appears to be little overlap between isolates, meaning that the diversity of encoded *Pf*EMP1 is vast. Indeed, in one study over 8000 variants were sequenced, yet the authors failed to detect significant overlap between parasite lines (Barry et al., 2007).

Other multi-gene families also exist in these locations, often in tandem with *var* genes, including the: *rif* (Repetitive Interspersed Family), *stevor* (Sub-Telomeric Variable Open Reading Frame), *Pfmc-2tm* (*P. falciparum* Maurers Cleft 2 TM) and *surf* genes.

The *rif* genes form the largest family in *P. falciparum*, containing approximately 150 copies per haploid genome (Gardner et al., 2002). The genes consist of a short 5' exon and a second exon, which encodes the majority of the RIFIN protein. Unlike the *var* genes, the *rif* repertoire was fairly consistent between the three sequenced *P. falciparum* isolates 3D7, HB3 and IT4 (Bultrini et al., 2009). The authors' explanation for this was that whilst stochastic events may ensure the generation of novel variants, substantial homogenization prevents the loss of function in the encoded proteins.

The RIFIN proteins, of 30-45kDa, were initially identified by iRBC surface protein radio-labelling and SDS PAGE (Cheng et al., 1998, Weber, 1988). Two putative trans-membrane (TM) domains were predicted (Gardner et al., 1998), and the region between was the most variable, containing many hydrophobic amino acids, and was predicted to form an extracellular 'loop' (Cheng et al., 1998). The *rif, stevor* and *Pf*MC2-TM gene families share a similar predicted protein structure and export motifs, so they have been proposed to be a super-family (Dzikowski et al., 2006, Sam-Yellowe et al., 2004). The 33 member- *stevor* gene family encodes products of approximately 30 kDa (Cheng et al., 1998), whilst the *Pf*MC2-TM family encodes 11 proteins, of approximately 27 kDa (Sam-Yellowe et al., 2004).

Another small multi-gene family exists in *P. falciparum*, comprising the 10 *surf* genes, of which three are predicted pseudogenes (Winter et al., 2005). These genes, like the other VSA families appear to have no close orthologues in other *Plasmodium* species.

Conversely, the *Plasmodium* interspersed repeat (*pir*) genes are present in many *Plasmodium* species but not *P. falciparum*, and were first described in *P. vivax*, where they were named *vir* (Del Portillo et al., 2001). The name was coined as an analogy for *P. vivax* variant genes, upon the discovery of 32 novel genes within a chromosome end from a *P. vivax* field isolate (Del Portillo et al., 2001). Subsequently orthologues were also identified in primate (*P. knowlesi, kir.* (Pain et al., 2008)) and rodent malaria species (*P. berghei, bir; P. chabaudi, cir; P. yoelii, yir.* (Janssen et al., 2002)). These genes form the largest multi-gene family within each genome, with the exception of *P. knowlesi*, encoding up to 5% of the total genes.

pir genes are typically found in sub-telomeric locations (*vir, bir, cir* and *yir*), although non-sub-telomeric *yir* family members are also found (Fonager et al., 2007). The *kir* genes however, are dispersed throughout the chromosomes, associated with interstitial telomeric repeat sequences (Pain et al., 2008), suggesting that the *kir* distribution was

created by recombination (Lin and Yan, 2008).

Most *pir* genes share the structure identified by del Portillo and colleagues, comprising a short first exon, long second exon and a highly conserved final exon (Del Portillo et al., 2001). PIR proteins contain most variability in their N-terminal sequence, although conserved cysteine residues exist, followed by a predicted trans-membrane domain (Del Portillo et al., 2001, Janssen et al., 2002).

1.8 *P. falciparum* variant surface antigens

In accordance with the long-held view that that *P. falciparum* infection alone caused severe malaria syndromes (described above), efforts to characterize and investigate the function of VSAs have predominantly concentrated on the *Pf*EMP1, RIFIN and STEVOR proteins. These efforts have focussed on particular aspects to elucidate possible functions, namely the proteins' cellular localization, regulation, adhesive properties and immune recognition, which will be discussed below.

1.8.1 VSA protein export

As indicated by the name 'variant surface antigens', most of the VSA families in *P. falciparum* have been located at or near the surface of iRBCs. Transport of these proteins from the developing parasite to the iRBC surface requires translocation across three plasma membranes, the parasite membrane, the parasitophorous vacuole membrane and the iRBC membrane. This will be briefly described before the localizations of *P. falciparum* VSAs are discussed.

Wickham and colleagues have shown that a recessed amino-terminal endoplasmic reticulum (ER) signal sequence is sufficient to cross the parasite plasma membrane, but not beyond (Wickham et al., 2001). In addition, the *Plasmodium* export element (PEXEL) or vacoular transport signal (VTS) sequence: RXLXE/Q/D is found in many exported proteins, approximately 35 amino acids downstream of the signal sequence (Marti et al., 2004, Hiller et al., 2004). This sequence alone has been shown to be sufficient for protein export in *P. falciparum* [for example: (Bhattacharjee et al., 2006)], hence several groups have searched for the presence of a PEXEL motif to predict the malaria 'secretome' (Sargeant et al., 2006, Van Ooij et al., 2008). The PEXEL motif is also cleaved in the ER, creating a new N-terminal sequence beginning with XE/Q/D, with an acetyl group at the terminal residue (Chang et al., 2008, Boddey et al., 2009).

Following PEXEL-cleavage, protein export is mediated, at least in part, by the novel '*Plasmodium* translocon for protein export' (PTEX) machinery identified by de Koning-Ward and colleagues (De Koning-Ward et al., 2009). The authors hypothesize that proteins are recognized by members of the PTEX complex as they reach the vacuolar

space, whereupon the HSP100/ClpA/B chaperone-related protein (HSP101) mediates unfolding, and the protein passes through a central pore, most probably provided by 'exported protein-2' (EXP2). On the cytosolic face of the PVM, host and / or other exported parasite proteins may contribute to protein refolding (De Koning-Ward et al., 2009, Gehde et al., 2009).

Not all exported proteins appear to contain a PEXEL motif however (Blisnick et al., 2000, Spycher et al., 2003, Hawthorne et al., 2004). It could be that these 'PEXEL-negative exported proteins' (PNEPs) actually contain PEXEL-equivalent motifs, that there are other properties of the N-terminus which are required for export across the PVM using the PTEX translocon, or that alternative export mechanisms are involved [reviewed by (Spielmann and Gilberger, 2010)].

Once across the PVM, proteins bound for export probably enter a secretory vesicular pathway to reach the iRBC membrane, which may include the membranous structures known as Maurer's Clefts (Hanssen et al., 2008). The possession of a TM domain appears to aid this process (Spycher et al., 2006, Saridaki et al., 2009).

1.8.2 Cellular localization

*Pf*EMP1 proteins bind to the electron-dense conical knobs of the iRBC membrane via KAHRP (Aikawa et al., 1985, Waller et al., 1999). *Pf*EMP1 proteins were identified following their detection at the surface of iRBC by cell-surface radio-iodination experiments (Howard et al., 1988). RIFINs were initially detected at the iRBC surface, appearing to be transit via the Maurer's clefts (MC) in association with *Pf*EMP1 (Fernandez et al., 1999, Kyes et al., 1999, Haeggström et al., 2004).

Two sub-types of RIFINs have since been identified (Joannin et al., 2008), where approximately two thirds of the RIFINs contain a 25 amino acid insertion sequence (Gardner et al., 2002, Pizzi and Frontali, 2001), located approximately 66 amino acids downstream from the PEXEL protein export motif, (Marti et al., 2004). RIFINs containing this sequence were designated type A (Joannin et al., 2008) and have been shown to display different cellular localizations to members of the other sub-family, B. In merozoites, A type RIFINs were located at the apical end, next to AMA1, whilst B-types were more dispersed throughout the cell, but not in the uniform distribution of MSP1 (Petter et al., 2007), Figure 10. Petter and colleagues showed that in ring,

trophozoite and schizont stages, A type RIFINs were both associated with the parasite and transported to the iRBC surface via MC, whilst the B type were retained within the parasite (Petter et al., 2007), Figure 10. A study using GFP-tagged RIFINs confirmed the localization pattern of B-types and suggested that the C-terminus was a crucial determinant for whether they were exported beyond the PV (Sijwali and Rosenthal, 2010).

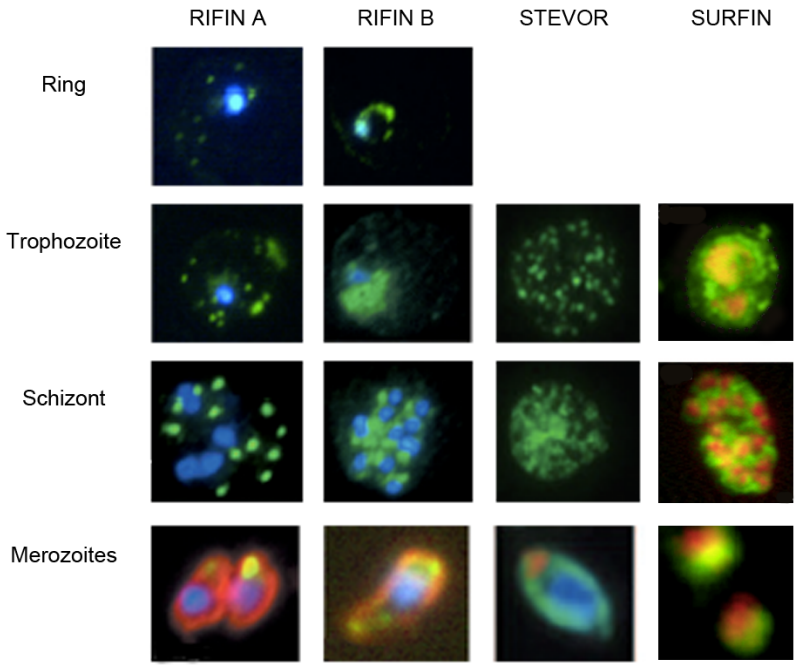


Figure 10: Cellular localization of VSA family members in different stages of *P. falciparum* intra-erythrocytic development.

Adapted from (Petter et al., 2007, Blythe et al., 2008, Winter et al., 2005). *P. falciparum* parasites of ring, trophozoite, schizont and merozoite stages are shown, stained with Abs generated against proteins from the VSA families RIFIN A- and B- types, STEVOR and SURFIN 4.1, in green. Parasite nuclei were stained with Hoechst 33342, blue (RIFIN and STEVOR), or propidium iodide, red (SURFIN). In addition, merozoite localization of VSAs was compared to other proteins, shown in red: MSP1, for RIFINs, and AMA1, for STEVOR.

STEVOR proteins and transcripts have also been detected in many stages of the *Plasmodium* lifecycle: in merozoites, in gametocytes and sporozoites (Florens et al., 2002, Florens et al., 2004, Le Roch et al., 2003, Le Roch et al., 2004, Bozdech et al., 2003b, Lavazec et al., 2006, McRobert et al., 2004, Khattab et al., 2008, Khattab and Meri, 2011). In trophozoite and schizont stage parasites, STEVORs have been observed both at the iRBC surface and in the MC, during trafficking (Kaviratne et al., 2002, McRobert et al., 2004, Przyborski et al., 2005, Blythe et al., 2008, Niang et al., 2009,

Khattab and Meri, 2011). Whilst STEVOR proteins display different localizations to the RIFINs in trophozoite stages, the apical expression observed in merozoites is similar to A-type RIFINs (Blythe et al., 2008), Figure 10.

Recent bio-informatic analyses suggest that in fact the two-TM domain gene model is wrong for A-type RIFINs and only one TM domain exists (Bultrini et al., 2009). The authors also predicted that these proteins have an 'armadillo-like' structural fold, which is found in diverse circumstances, facilitating protein-protein interactions (Coates, 2003). In line with the suggestions that the 2-TM model may not best represent the RIFIN repertoire (Bultrini et al., 2009), evidence has emerged that the majority of the STEVOR protein is also exposed on the iRBC surface, likely anchored by the C-terminal TM domain (Niang et al., 2009). Whilst this supports the similarities previously observed between these multi-gene families, it remains to be seen whether these proteins truly form a super-family, performing similar functions.

Within erythrocytic stages, *Pf*MC2-TM family members have been localized to the MC and in the knob structures of iRBC (Lavazec et al., 2006), but iRBC surface localization has not been proven. The inter-TM loop in these proteins is small, only three amino acids (Sam-Yellowe et al., 2004) and this region may be exposed on the iRBC surface, as proposed by Sam-Yellowe and colleagues.

Finally, members of the SURFIN family have been located at the surface and the PV of iRBCs and also in merozoites (Winter et al., 2005). Interestingly, as observed for members of the RIFIN family, different SURFIN variants appear to have different cellular localizations, for example SURFIN 4.1 was detected forming an 'amorphous cap' at the apical end of merozoites, whilst the distribution of SURFIN 4.2 was more discrete in the centre of the merozoite cell (Winter et al., 2005). Similarly to other VSA families, their function(s) are as yet unknown.

1.8.3 Regulation of expression

There has been a great deal of interest in understanding how *var*/*Pf*EMP1 expression is regulated following early data suggesting that only one *Pf*EMP1 is present on the surface of an iRBC (Biggs et al., 1991, Scherf et al., 1998, Roberts et al., 1992, Biggs et al., 1992, Chen et al., 1998). It seems that ring stage parasites may transcribe several *var* genes, but mutually exclusive expression is achieved during mature stages of

development (Duraisingh et al., 2005, Dzikowski et al., 2006, Deitsch et al., 2001), whereby the dominant transcript is the only one translated and exported to the iRBC surface (Chen et al., 1998).

The regulation of *var* transcription is highly complex, involving three layers of regulation [reviewed by (Scherf et al., 2008, Chookajorn et al., 2008, Dzikowski and Deitsch, 2009)]. The first layer comprises the cis-regulatory sequences of *var* gene promoters to which transcription factors bind, such as members of the APiAP2 family (Llinás et al., 2006). The intron also contains a promotor that drives expression of the intron sequence in most *var* genes (Epp et al., 2009). This appears to act in silencing and may contribute to *var* recognition by the mechanism regulating their mutually exclusive expression (Deitsch et al., 2001, Frank et al., 2006, Dzikowski et al., 2007). The second layer comprises epigenetic histone marks, such as methylated H3K27 and tri-methylated H3K9 (Duraisingh et al., 2005, Lopez-Rubio et al., 2007), which maintain *var* genes respectively in their active or silent state for several cell cycles (Horrocks et al., 2004a, Frank et al., 2007). The final layer is the sub-nuclear localization of chromatin, whereby chromatids appear to be re-positioned into distinct nuclear sites to allow active transcription (Ralph et al., 2005).

Importantly, the majority of iRBCs also express the same *Pf*EMP1 as the schizont from which they were derived (Roberts et al., 1992). In general this leads to expression of the same *Pf*EMP1 by the majority of iRBCs during *P. falciparum* infection (Marsh and Howard, 1986, Bull et al., 1998, Giha et al., 1999, Ofori et al., 2002), and the major variant may switch over a period of weeks (Staalsoe et al., 2002). Recent data obtained from modelling *var* transcription patterns during *in vitro* culture indicate that certain variants dominate in the parasite population due to a high likelihood of switching to these variants when other *var* genes are switched off in combination with these genes having slow switching-off rates (Recker et al., 2011).

After RBC invasion, the first VSA transcripts produced are *var*, followed by *rifin* and then *stevor* (Kyes et al., 2001, Kaviratne et al., 2002), whilst *Pf*MC2-TM genes were transcribed between the times of *rifin* and *stevor* genes (Sam-Yellowe et al., 2004). No co-expression with *var* genes has been observed either for *stevor* (Sharp et al., 2006), or *rifin* genes (Tham et al., 2007). This indicates that, despite their chromosomal proximity to *var* genes and the presence of similar promoter sequences, the *rifin* and *stevor* genes are independently regulated.

Between four and 10 *Pf*MC2-TM proteins have been detected from trophozoite stage samples of iRBC, indicating that at the population level, several of these genes were expressed at the same time (Sam-Yellowe et al., 2004). This has also been observed for the *stevor* genes, where several different transcripts were detected within individual iRBCs (Kaviratne et al., 2002). Recent data however, indicate that although several STEVORs may be expressed within iRBCs, only one is exposed at the surface of trophozoite stages (Khattab and Meri, 2011). This suggests that STEVORs may also mediate mutually exclusive expression at the iRBC surface like *Pf*EMP1 family members, a phenomenon which may extend to the other VSA families.

1.8.4 Cyto-adhesion: rosetting and sequestration

The severe disease syndromes of pregnancy associated malaria and cerebral malaria have been associated with cytoadhesion to the receptors CSA, HA, and ICAM1, respectively (Fried and Duffy, 1996, Newbold et al., 1999, Newbold et al., 1997, Ochola et al., 2011), whilst CD36 binding is a characteristic of many clinical isolates, and may be associated with less severe disease outcomes (Ockenhouse et al., 1992, Reeder et al., 1994, Ochola et al., 2011). At present it is not clear exactly how the phenomenon of iRBC sequestration occurs, although the process is better understood in the placenta, than for example, in cerebral malaria. *P. falciparum* adhesion to placenta has been shown to involve the semi-conserved VAR2CSA *Pf*EMP1 protein binding to CSA (Fried and Duffy, 1996, Salanti et al., 2004), present in the intervillous space of placental tissue. This blocks blood supply to the foetus, which is exacerbated by the infiltration of monocytes (Walter et al., 1982). The development of novel technologies, such as whole mouse imaging (Franke-Fayard et al., 2005), multi-photon imaging (Ortolano et al., 2009), and transcriptional profiling (Vignali et al., 2011) are likely to elucidate the mechanisms underlying parasite cyto-adhesion and cell:cell interactions.

*Pf*EMP1 proteins have the capacity to cyto-adhere via several host receptors: CD36 (Baruch et al., 1996, Robinson et al., 2003, Mo et al., 2008), ICAM 1, thrombospondin (Baruch et al., 1996), CSA (Fried and Duffy, 1996), and to form rosettes by binding to uninfected RBCs (Rowe et al., 1997, Chen et al., 1998). Certain binding phenotypes have been attributed to individual domains, for example CIDR1α domains from many parasite isolates bind to CD36 (Robinson et al., 2003), whilst the DBLβ domain and c2 region are both required for binding to ICAM-1 (Springer et al., 2004). It is possible

that inclusion of two or more of these domains in an individual *Pf*EMP1 molecule could allow binding to several host receptors, a feature that has been associated with severe disease (McCormick et al., 1997, Ockenhouse et al., 1992, Heddini et al., 2001) and also observed using selected *in vitro* lines (Smith et al., 2000, Chen et al., 2000).

Due to their different domain compositions, *Pf*EMP1 sub-groups have different binding capacities. Most members of the Ups B and C groups bind to CD36 (Robinson et al., 2003), whilst the Ups A proteins do not, and may instead bind to a range of receptors (Kyes et al., 2007, Janes et al., 2011). It is thought that the variants associated with severe disease preferentially bind host receptors in, for example, brain endothelium. This would explain observations of iRBC accumulation in brain and heart tissue from post-mortem analysis of malaria fatalities (Montgomery et al., 2007). Rosetting also appears to contribute to severe disease since sera from children experiencing cerebral malaria were unable to disrupt rosettes, unlike sera from children experiencing uncomplicated infection (Carlson et al., 1990, Rowe et al., 1995, Horata et al., 2009).

It has been suggested that variants appear during infection in a pre-determined pattern (Bull et al., 1998, Bull et al., 2000, Nielsen et al., 2002, Horrocks et al., 2004b). Although this is difficult to determine in the field (Bull et al., 2005a, Bull et al., 2005b), evidence from experimental human infections support this idea. Lavstsen and colleagues studied naïve adults infected with *P. falciparum* sporozoites, observing that parasites emerging into the bloodstream from the liver expressed an assortment of variants, but after only two or three cell cycles the majority of transcribed *var* genes belonged to the Ups A group (Lavstsen et al., 2005). The recent data from Recker and colleagues also supports the idea of pre-determined *var* expression, in the absence of immune selection, such that infection of naïve individuals would be more likely to result in the expression of a narrow range of *var* genes (Recker et al., 2011).

The functions of RIFIN proteins are unknown. They were initially thought to mediate rosetting (Helmby et al., 1993, Kyes et al., 1999, Fernandez et al., 1999), although whether the RIFINs may contribute to rosette formation is unclear as *Pf*EMP1 family members are now known to bind uninfected RBCs (Rowe et al., 2000, Russell et al., 2005). Other evidence for RIFIN-mediated cyto-adhesion is lacking, although one study found identical *rif* transcription in parasite lines displaying different binding characteristics (Cabral and Wunderlich, 2009), suggesting that iRBC adhesion, at least to CD36, ICAM-1 and Selectin does not require RIFIN proteins. Conversely, differences in *stevor* transcription have been observed after selection of parasite lines

for adhesion to the receptors CD36 and CSA (Khattab et al., 2008), suggesting that these proteins may contribute to iRBC sequestration.

1.8.5 Parasite survival *in vivo*

It has been shown that parasite isolates down-regulate the transcription of a number of VSA genes upon adaptation to *in vitro* culture: *var*, *rif* and *stevor* (Mackinnon et al., 2009). Similar down-regulation was also observed in a splenectomised patient experiencing severe malaria, with specific down-regulation of *var*, *stevor* and A-type *rif* but not B-type *rif* or *Pf*MC 2TM genes (Bachmann et al., 2009). This indicates that *Pf*EMP1, STEVOR and the A-type RIFIN proteins are not essential for parasite survival in the absence of splenic selection. Interestingly, the B-type RIFINs were not down-regulated during asplenic infection (Bachmann et al., 2009), which probably indicates that these proteins are not exposed to selection, due to their PVM localization within iRBCs (Petter et al., 2007). In fact, early phylogenetic analysis of the RIFINs found evidence of different selective pressures applying to different subsets, which the authors suggested could be due to different exposure to the immune system (McInerney et al., 2003). Following identification of the two subfamilies, Joannin and colleagues indicated that sub-functionalization has occurred within the RIFINs (Joannin et al., 2008), namely that the sequences of RIFINs belonging to the two sub-groups have significantly diverged, and the proteins may perform distinct functions. This is supported by their different cellular localizations [Figure 10, (Petter et al., 2007)].

Similarly, to the B-type RIFINs, *Pf*MC 2TM proteins did not appear to be subject to splenic selection (Bachmann et al., 2009). Due to their MC localization within iRBC, these proteins have been hypothesized to involve trafficking of other VSA to the iRBC surface (Sam-Yellowe et al., 2004). If this were true, the *Pf*MC 2TM and possibly B-type RIFINs could mediate transport of those VSAs which appear essential to *in vivo* parasite survival: the *Pf*EMP1, A-type RIFIN and STEVOR proteins.

The functions of these proteins could be sequestration to evade splenic clearance of iRBCs, as discussed above. Another explanation could be that these proteins are required for evasion of immune selection, which would be severely limited in the absence of the spleen, a major lymphatic organ. In the absence of splenic selection, it is feasible that iRBC lacking exposed VSAs could proliferate more quickly than iRBC

expending energy to produce and traffic these proteins, and hence these parasites would come to dominate the infection.

1.8.6 Immune evasion

Although the functions of the different VSA families remain to be fully determined, it is evident that their variant expression acts to preserve expression of at least one member in the face of specific immune responses to other family members. The presence of Abs to different variants results in differential *in vivo* selection pressures within the parasite population. Abs recognizing VSAs are likely to play an important role in mediating parasite clearance, performing functions such as opsonization and complement mediated lysis of iRBC (Beeson et al., 2008). In addition, Abs generated against certain VSA, notably *Pf*EMP1, are likely to inhibit cyto-adhesion.

During infection, it is thought that sequestration provides a selective advantage for iRBC to escape splenic clearance. Thus in the absence of Abs, parasites expressing strong adhesive molecules will proliferate to dominate the infection. Furthermore, even dead iRBCs can continue to sequester (Hughes et al., 2010), meaning that vascular occlusion and other neurological phenomena could continue to occur after anti-malarial treatment.

An important consideration for our understanding of severe malaria is how the generation of specific Abs to the *Pf*EMP1 domains mediating adhesion may influence parasite behaviour. The presence of Abs recognizing the dominant, strongly adhesive domains means that iRBCs expressing these proteins would no longer be able to sequester, and other *Pf*EMP1 variant-expressing iRBCs would have a selective growth advantage during infection (Hviid, 2010, Marsh and Howard, 1986, Giha et al., 1999, Ofori et al., 2002). The resulting antigenic variation may be explained by two different theories:

- Phiri and colleagues have proposed that parasites expressing more effective adhesion molecules will out-compete other iRBCs to bind available host receptors, particularly within the currents of the bloodstream. This has been demonstrated under flow conditions in the absence of immune pressure (Phiri et al., 2009). Upon acquisition of immunity to the most adhesive variants, other

variants would experience a selective advantage, which would enable iRBC proliferation, and these variants would become dominant.

- An alternative scenario is that different domains arrangements between *Pf*EMP1 variants means that Abs recognizing, for example, the DBL1α domain, will be likely to cross-react with other variants that also contain this domain, or similar epitopes. In this case, any *Pf*EMP1 molecules containing domains to which cross-reactive responses were induced would be selected against, allowing the proliferation of iRBCs expressing other variants. The generation of Abs cross reactive with many *Pf*EMP1 domains or variants have been described, both from experimental infection (Elliott et al., 2007), and in the field (Mackintosh et al., 2008, Chattopadhyay et al., 2003). Interestingly, *Pf*EMP1 variants belonging to the Ups A group have been shown to contain more cross-reactive Ab epitopes than variants belonging to the other sub-groups (Joergensen et al., 2006).

Mathematical modelling has been able to reproduce both of these scenarios: Recker and colleagues suggested that cross reactive immune responses could effectively structure the *Pf*EMP1 expression patterns observed during infection (Recker et al., 2004); whilst van Noort and collegues reproduced the association of different variants with severe and uncomplicated malaria, by combining adhesive phenotypes with cumulative Ab repertoires (Van Noort et al., 2010). Both mechanisms could play equally important roles in structuring the antigenic variation of *Pf*EMP1 during infection.

Evidence is increasing that other VSAs are also involved in the process of antigenic variation or immune evasion. The *Pf*MC2-TM genes, along with the *stevor*s, have also been shown to undergo switching of expression at rates similar to the *var* genes (Lavazec et al., 2007). Transcription of *stevor* genes has been detected in up to 90% of iRBC in fresh parasite isolates, which decreased upon adaptation to *in vitro* culture to only 10% of iRBC (Blythe et al., 2008), similar to that seen for the *var* and *rif* genes (Mackinnon et al., 2009). Niang and colleagues have also confirmed the iRBC surface exposure of STEVORs, and that changes in their expression profile directly correlated with the immunogenic properties of the iRBC (Niang et al., 2009).

1.8.7 Acquisition of immunity to *P. falciparum* VSAs

Although the potential number of VSA variants is immense, they are targets of protective immunity (Marsh et al., 1989, Bull et al., 1998, Dodoo et al., 2001, Ofori et al., 2002, Kinyanjui et al., 2004a, Yone et al., 2005, Magistrado et al., 2007), and it seems that acquisition of immunity to a small fraction is associated with protection against severe disease (Bull et al., 2005b, Kaestli et al., 2006).

Clinical malaria has been linked to the Ups A and B/A *Pf*EMP1 groups, whilst asymptomatic infections have been associated with parasites expressing Ups B and C types (Jensen et al., 2004, Kaestli et al., 2006, Kyriacou et al., 2006, Rottmann et al., 2006). All group A and some of the B/A *Pf*EMP1 variants contain two cysteine residues within the DBLα domain, termed 'cys 2' (Bull et al., 2005a, Bull et al., 2007). These variants have been detected during infection of immuno-deficient children and in cases of cerebral malaria (Bull et al., 2005a, Kyriacou et al., 2006, Warimwe et al., 2009), appearing to be expressed in the face of a poorly developed Ab response.

The humoral response to *Pf*EMP1 domains has been investigated by several groups, finding variable responses towards different domains that tended to increase with age (Mackintosh et al., 2008, Vestergaard et al., 2008). Cham and colleagues have definitively shown that children acquire Abs to Ups A and B/A *Pf*EMP1 variants first (Cham et al., 2009, Cham et al., 2010). In areas of high transmission, the Ab repertoire increased more quickly, but Abs were acquired to DBL-like domains in a similar order regardless of endemicity. Once Abs against the A/B type were acquired, a broader range of *Pf*EMP1 variants was encountered, which took longer to generate effective immunity against, (Figure 8). This ordered acquisition of Abs neatly explains the three phases of *Plasmodium* infection that is observed in endemic areas, described in section 1.4.

Whilst immunity to *P. falciparum* must comprise variant specific responses, including the UpsA variants of *Pf*EMP1, it is likely that cross-reactive immune responses also occur, which are able to overcome antigenic variation to some degree. Indeed there have been reports of boosting responses during heterologous infection (Kinyanjui et al., 2004a, Bull and Marsh, 2002, Ofori et al., 2002). In addition, certain *Pf*EMP1 domains also appear capable of inducing cross-reactive immune responses (Gamain et al., 2001, Mackintosh et al., 2008). This finding, in concert with the association of only a small proportion of the *Pf*EMP1 repertoire and severe disease, has led to proposals for development of *Pf*EMP1-based vaccines (Chen, 2007).

Vaccine design would need to consider which epitopes may stimulate protective responses, particularly of T helper cells, which coordinate adaptive immunity. *Pf*EMP1 epitopes that elicit cross-reactive CD4 responses would thus be more likely to contribute to protection by stimulating the proliferation of both cross-reactive and variant specific B-cells. The cellular response to *Pf*EMP1 has been investigated using the three most conserved domains: the extracellular DBLa, and CIDRa domains and a region from the intracellular region within exon 2 (Allsopp et al., 2002, Sanni et al., 2002, Das et al., 2007, Ndungu et al., 2006). Surprisingly, CD4 T cells from both malaria-naïve and -exposed individuals had measurable proliferative and cytokine responses to the CIDR domain, suggesting that this region may contain epitopes similar to non-*Plasmodium* proteins to which the donors had all been previously exposed (Ndungu et al., 2006, Allsopp et al., 2002). Specific responses were observed to the other domains, suggesting that these regions contain T cell epitopes that may be useful vaccine components (Allsopp et al., 2002). In particular the peptide SDITSSESEYEEMDINDIYVPGS from the exon 2 region has been shown to induce more consistent magnitudes of proliferative, IL4 and IFNγ responses than several other peptides from this region, and higher proliferative and IL4 responses in immune adults than malaria patients [in Delhi, India (Das et al., 2007)].

In fact, vaccine development is underway for the unusual, conserved *Pf*EMP1 variant VAR2CSA (http://cmp.ku.dk/english/research/teamvar2csa/), found in all *P. falciparum* genomes, which mediates sequestration of iRBC to the placenta during PAM (Fried and Duffy, 1996). Immunity to PAM specifically correlates with the presence of Abs to conserved regions of VAR2CSA (Rogerson et al., 2007), to which cross-reactive Abs may be generated by immunization (Avril et al., 2010).

Less is known about immunity to the non-*Pf*EMP1 VSA families in *P. falciparum*. RIFIN and STEVOR proteins are recognized by Abs present in immune serum from clinically immune adults (Fernandez et al., 1999, Leech et al., 1984, Schreiber et al., 2008, Abdel-Latif et al., 2003, Abdel-Latif et al., 2002). The presence of anti-STEVOR Abs in Ghanian infants has been shown to correlate with the number of parasitaemia episodes experienced, and over time the children developed a pattern of high, transient or low humoral responses (Schreiber et al., 2008). It is not clear whether these Abs contribute to protection from clinical disease, since several studies appear to show either no association of anti-RIFIN or STEVOR Abs with protection (Abdel-Latif et al., 2003, Schreiber et al., 2008) or even that the presence of these Abs could be risk factors

for development of higher peripheral parasitaemia or cerebral malaria (Schreiber et al., 2006, Schreiber et al., 2008). In the latter study, the IgG2 and 4 isotypes of RIFIN specific Abs were exclusively detected in cerebral malaria cases, the authors suggesting that these Abs may play pathogenic roles during infection. An alternative explanation is simply that the acquisition of Abs against this vast number of VSA reflects the number of infections experienced, which are more likely to result in severe disease in young children.

Given the long-standing dogma that *P. falciparum* infection alone caused severe malaria syndromes (described in Table 1), it is not surprising that the majority of studies investigating the generation of immunity to VSA have concentrated on those found in this species. It is likely that similar phenomena could also occur in other *Plasmodium* infections, which are not as benign as once thought.

1. 9 The *pir* / PIR 'super-family'

Non-*P. falciparum* species also undergo antigenic variation, mediate rosetting, and cytoadherance via receptors including CD36 (Mackinnon et al., 2002, Mota et al., 2000, Franke-Fayard et al., 2005, Carvalho et al., 2010, Brown and Brown, 1965, Gilks et al., 1990, McLean et al., 1982), so these phenomena must be mediated by VSAs besides *Pf*EMP1. The *pir* multi-gene family may be able to perform some of these functions, which has members in the majority of *Plasmodium* species described to date.

1.9.1 Classification of the PIR superfamily

The approach that has been widely utilized in description of multi-gene families, as observed earlier for the *Pf*EMP1 and RIFIN genes, is their classification into smaller sub-groups, based on shared characteristics. This has also been used to define the *vir* genes, initially describing 6 clusters: A-F (Del Portillo et al., 2001), and members of these sub-families were detected in parasites isolated from 32 patients (Merino et al., 2006). On completion of the *P. vivax* Salvador I genome sequence (Carlton et al., 2008), further sub-families were identified, giving a total of 12 sub-families from 76% of VIR amino acid sequences. The *yir* repertoire has been analyzed in a similar way, with 56% of *yir* genes clustering to form five sub-groups (Fonager et al., 2007). Orthologous members of these sub-groups were also identified in the *bir* repertoire, reflecting the recent evolutionary divergence between *P. yoelii* and *P. berghei* (Hayakawa et al., 2008).

A comparison of PIR sequences from the whole 'super-family' was carried out by Janssen and colleagues using approximately 8.5% of the total repertoire [157 sequences analyzed from an estimated 1843 total sequences, (Janssen et al., 2004)]. The relationships between 136 PIRs have been confirmed using different bio-informatic methodology (Cunningham et al., 2010), namely that PIR sequences displayed most similarity with their orthologues in the evolutionarily closest *Plasmodium* species. These data are summarized in Figure 11.

RIFIN and STEVOR sequences were also included in the analyses of Janssen and colleagues, with the suggestion that these may be distantly related to the PIR super-family (Janssen et al., 2004). As a consequence of the evolutionary distance between *P.*

falciparum and other *Plasmodium* species, these sequences did not cluster with the PIR sequences. The authors postulated that these families may share ancestral sequences, and noted that there was significant orthology between the intron of *rif* genes and the second *pir* intron (Janssen et al., 2004). The inclusion of the *rif, stevor* and *PfMC-2TM* genes within the *pir* super-family was based upon similarities in size, and conserved sequences present in the first intron, the authors proposing an ancestral relationship rather than necessarily conserved functions between these gene families. However, this inclusion has proved controversial [for example: (Joannin, 2010)], especially due to the divergence in amino acid structure between the *P. falciparum* gene families and the other *pir*.

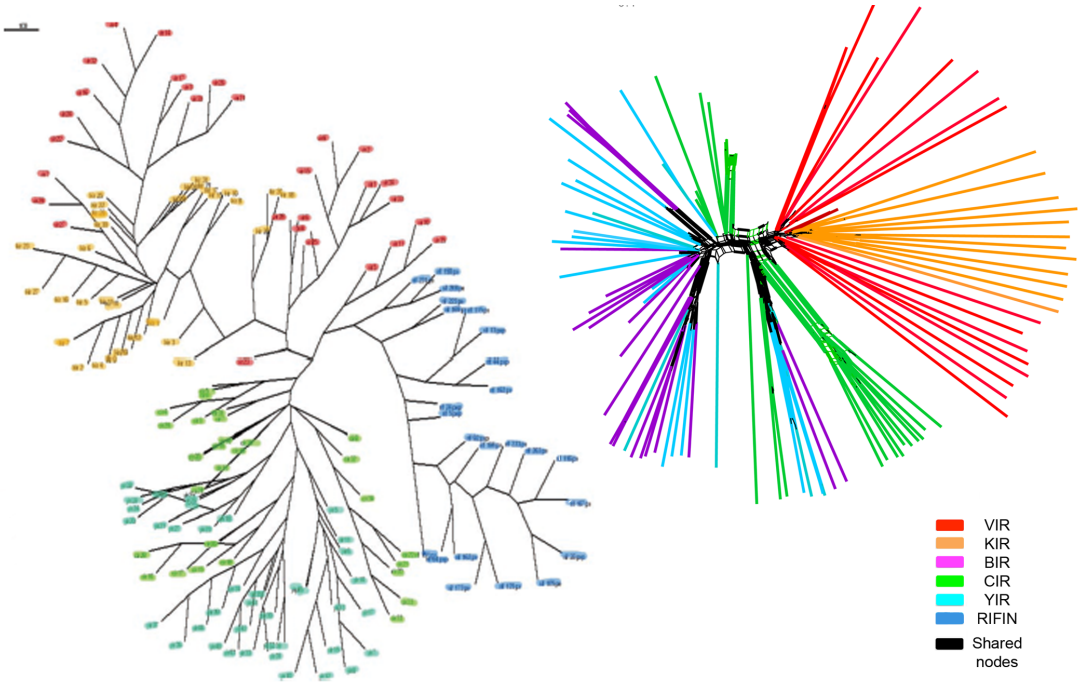


Figure 11: Sequence comparison between PIR super-family members.

Adapted from (Cunningham et al., 2010) (Janssen et al., 2004). The left tree was created using the maximum parsimony algorithm, and shows the relationships between 157 PIR and RIFIN sequences (Janssen et al., 2004), whilst the right hand image is a NeighborNet network (Bryant and Moulton, 2004) which shows the relationships between 136 PIR sequences (Cunningham et al., 2010) .

Indeed, since the functions of none of these multi-gene families have been elucidated, it could be argued that definition of any gene family purely based on their sequence characteristics risks being obstructive. Future experimental evidence showing

disparities or parallels between members of the variant antigen families should aid their definitive classifications.

1.9.2 *pir* / PIR expression

Northern blot analysis of *cir* transcription during the erythrocytic growth cycle indicated that late-trophozoite stages have most active transcription, which is reduced during schizogony (Janssen et al., 2002). Microarray studies carried out subsequently in *P. vivax* and *P. yoelii* suggest that at the population level, the pattern of *pir* transcription is more complex with many changes occurring (Bozdech et al., 2008, Cunningham et al., 2009). In these studies, 42% and 59% of the *yir* and *vir* repertoires were transcribed respectively.

Initial reverse-transcription PCR experiments have suggested that *yir* sub-families may be differentially expressed in different parasite stages (Fonager et al., 2007), however, this was not confirmed by a subsequent microarray investigation (Cunningham et al., 2009). Instead, a large proportion of *yir* were found transcribed at low levels, each gene having different transcriptional peaks at different stages of intra-erythrocytic development. This was the case even within cloned parasites, in the absence of lymphoid selection, where the expressed *yir* repertoire was substantially altered from the infecting parasite within 10–12 growth cycles (Cunningham et al., 2009).

Similarly to the *yir* genes, there was also no evidence that *vir* transcription correlated with the previously identified phylogenetic groups (Bozdech et al., 2008, Del Portillo et al., 2001, Carlton et al., 2008). However, two non-overlapping sets of *vir* transcription were detected in different stages of *P. vivax* development: shortly after RBC invasion and in schizonts (Bozdech et al., 2008). The waves of *vir* transcription were associated with a similar pattern for the *P. vivax* tryptophan rich antigen genes (*Pvtrag*), which have also been postulated to be involved in immune evasion (Jalah et al., 2005).

pir transcription has recently been compared between the erythrocytic stages of infected patients and their subsequent sexual stages (gametes / zygotes and ookinetes), generated by *in vitro* culture (Westenberger et al., 2010). The majority of *vir* genes showed lower expression than other genes at all stages, and strong differential regulation was only displayed by 9.9% of *vir*s (Westenberger et al., 2010).

Using *P. yoelii*, it has been possible to conduct controlled experiments to investigate *pir* transcription in pre-erythrocytic stages (Williams and Azad, 2010).The transcriptome was compared between WT sporozoites, radiation-attenuated sporozoites and liver stages. Differences in *yir* transcription were observed in that wild-type sporozoites retained a higher number of expressed *yir* genes than radiation-attenuated sporozoites, and that distinct groups of *yir* genes were expressed by sporozoites and 48hr liver stages, and 24hr liver stages contained an intermediate *yir* profile (Williams and Azad, 2010).

The detection of *yir* transcripts in sporozoites supports earlier proteomic data which found BIR proteins expressed at different stages of the parasite lifecycle (Hall et al., 2005). In this study almost 20% of the possible BIR repertoire was detected, with the vast majority expressed within specific, non-overlapping stages, and almost a tenth of BIR expressed within ookinete stage parasites (Hall et al., 2005). Together, these data suggest that PIR subsets may perform distinct roles within different parasite stages.

1.9.3 Regulation of *pir* expression

Clearly, *pir* genes are subject to differential gene regulation. To decipher the level of regulation applying to individual iRBCs, micro-manipulation has been used to separate single cells and investigate their *pir* transcription (Fernandez-Becerra et al., 2005, Cunningham et al., 2009). Surprisingly, considering the large proportion of *yir* and *vir* genes that were transcribed in the total blood population, only 1 - 3 *yir* were transcribed in a single iRBC. This suggests that whilst mutually exclusive *pir* transcription does not occur, transcription of these genes is under tight control. It is possible that not all transcribed *pir* are translated so mutually exclusive protein expression cannot yet be ruled out.

Very little is known about regulatory factors affecting *pir* genes, although an informatic analysis has identified a motif within the *yir, bir* and *cir* promoter regions, which may provide general regulatory control rather than differential regulation (Fonager et al., 2007). This study also identified alternative splicing events occurring during *yir* mRNA processing. Alternative splicing may alter the character of the translated YIR or contribute to regulation of expression by blocking translation of full-length *yir* and perhaps regulation upon transcription of other *yir*s, as postulated for other VSA families

(Al-Khedery et al., 1999, Taylor et al., 2000a).

1.9.4 PIR protein characteristics

In general PIRs do not appear to contain strong homology to proteins whose structures have been solved. Predictions of their secondary structure have identified multiple alpha helices (7-8) ranging from 12 to 20 amino acids in length, separated by coiled-coil regions (Janssen et al., 2004). Secondary structure predictions for the VIR and KIR repertoires reflect their diverse natures, containing more beta-stranded regions (Janssen et al., 2004).

Domains found in other VSA, such as the DBL and CIDR domains of *Pf*EMP1 members are absent in PIR proteins. Instead, PIRs are comprised of up to three as yet uncharacterized PIR domains followed by a trans-membrane region (Pain et al., 2008). Their predicted arrangements are shown in Figure 12. Although the majority of PIRs do contain a predicted TM domain (Janssen et al., 2004, Pain et al., 2008), this is only predicted for half of the VIR proteins (Carlton et al., 2008), which may indicate that half of the proteins perform other functions to the other VIRs.

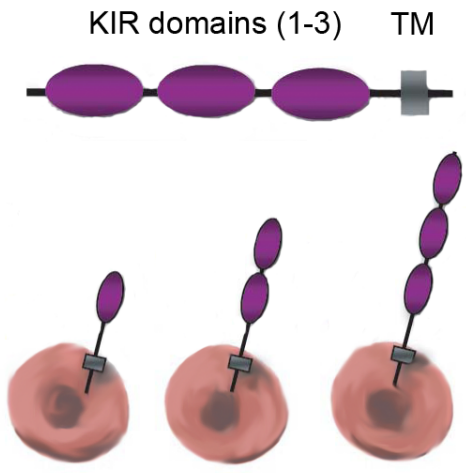


Figure 12: Domain organization predicted for the KIR proteins.

Adapted from (Pain et al., 2008). The three predicted KIR domains are shown in purple, followed by the TM domain, top panel. KIR proteins may contain one, two or three domains, which are thought to be exposed at the surface of iRBC, as indicated in the bottom panel.

Some PIR proteins also contain sequences similar to the PEXEL /VTS motif, associated with protein export to the iRBC surface in *P. falciparum* (Hiller et al., 2004, Marti et al., 2004). This is found in 46% of VIR in the Sal I strain, and in 69% of VIR sequenced from 32 patient isolates, 11% of YIR, but not in any KIR sequences (Pain et al., 2008, Fonager et al., 2007, Merino et al., 2006, Carlton et al., 2008). Instead, a different, potential export sequence has been identified in three quarters of the KIR repertoire: ZLPS [where Z encodes a hydrophilic residue, (Pain et al., 2008)].

Several features of PIR proteins are shared with members of the RIFIN and STEVOR families, including the typical PIR alpha helical secondary structure. In fact, *rif* genes could be identified from the *P. falciparum* genome using hidden Markov models designed for detection of rodent *pir* genes [with low significance, (Janssen et al., 2004)]. In addition, certain VIR sub-families display similarities with other *P. falciparum* multi-gene families: sub-family A members contain a cysteine rich domain at their N-termini, like the SURFINs (Merino et al., 2006, Winter et al., 2005), whilst sub-family D members share features with *Pf*MC-2TM proteins including an N-terminus which is weakly hydrophobic, conserved cysteine residues and proline residues in one of the TM domains, and a lysine rich C-terminus, (Merino et al., 2006).

1.9.5 Function(s) of PIR proteins

The function(s) of PIRs are unknown, although certain features of the proteins may provide an insight into their possible roles. The cellular localization provides a crucial insight into protein function and several investigations have been carried out to elucidate the localization of PIR within erythrocytic stages.

Numerous methodologies have been used to show that at least some PIR reside at or near the iRBC surface: indirect immunofluorescence assay using both fixed and live iRBC, flow cytometry using live iRBC and proteome analysis of detergent-resistant membrane-associated proteins (Fernandez-Becerra et al., 2005, Janssen et al., 2004, Del Portillo et al., 2001, Cunningham et al., 2005, Di Girolamo et al., 2008). In the latter study, an endogenous *bir* gene was GFP-tagged, and maximal fluorescence intensity was detected in trophozoite stage parasites, confirming earlier transcriptional data which had suggested peak PIR expression occurred at this stage of development (Di Girolamo et al., 2008).

Surface localization of PIR proteins is perhaps surprising, since few of the proteins contain export motifs, as discussed above. However, reports of protein export in the absence of PEXEL/VTS sequences suggest that alternative export pathways exist, which may utilize an overall N-terminal negative charge, as observed for skeleton-binding protein 1 (Saridaki et al., 2009). The localization of some PIRs at the iRBC surface, and the detection of BIR family members within detergent resistant lipid rafts, could indicate that PIR perform environmental interactions such as signalling and protein trafficking, functions of other lipid-raft associated proteins (Di Girolamo et al., 2008).

Given the fact that a small fraction of PIRs are found at the iRBC surface, and that different PIRs are expressed throughout the erythrocytic cycle, it seems likely that different PIRs may have distinctive cellular localizations. It could be that particular subsets perform individual functions, and given the similarities of certain VIR sub-groups with SURFIN and *Pf*MC-2TM sequences, it is possible that these proteins may also be found located in merozoites and at the MCs (Winter et al., 2005, Sam-Yellowe et al., 2004).

As described above, the *pir* genes represent the largest multi-gene family in *P. vivax, P. yoelii, P. berghei* and *P. chabaudi.* Their function(s) must be crucial for parasite survival otherwise such redundancy would have been quickly lost by stabilizing selection. The presence of at least some PIR at the iRBC surface may well indicate that they are involved in host-interactions. Potential host-interactions of the PIR proteins are discussed below, including: antigenic variation, immune interactions and sequestration.

1.9.5 i) Antigenic variation

Requirements that must be fulfilled for a multi-gene family to truly mediate antigenic variation are stringent (Turner, 2002). The five major criteria proposed by Turner are presented here, in context with features of the *pir* family.

1) Variant antigens must have the capability to express several different antigens; a feature easily fulfilled by the large size of the *pir* family.

2) Variants should ideally be immuno-dominant over conserved antigens, each variant containing unique conformational determinants to which immune

responses are directed. The diversity of *pir* sequences accounts for this factor, VIR proteins sharing as little as 20% sequence identity (Carlton et al., 2008), although there is no evidence as yet that PIRs are immuno-dominant.

3) Sub-populations of iRBCs expressing different variants must be capable of growth to compensate for parasites killed by effective immune responses. Although this has not been definitively shown regarding PIR, iRBCs isolated from different peaks of *P. chabaudi* infection express distinct surface antigens (McLean et al., 1982a), which are likely to be encoded by the *cir* genes as these comprise the largest multi-gene family in *P. chabaudi* (Janssen et al., 2002).

4) The differential expression of variants must not affect the expression of other genes, for example being induced as part of a physiological response. *pir* transcription has been shown to change without associated differences in other genes, and in different patient isolates (Bozdech et al., 2008, Cunningham et al., 2009, Westenberger et al., 2010, Williams and Azad, 2010).

5) Mutually exclusive transcription should occur to prevent exposure of all variants to the immune response and thus exhausting the repertoire of variants. This is not the case for *pir* genes, however transcription of 1 - 3 *pir* genes per iRBC nonetheless demonstrates tight control of *pir* regulation (Cunningham et al., 2009), and is similar to the number of *var* genes that may be transcribed in ring stage *P. falciparum* iRBCs (Chen et al., 1998). Similarly, the parasite population must retain the capacity to present different variants in waves with relatively little overlap. This has been shown within the *vir* repertoire (Bozdech et al., 2008).

Despite non mutually exclusive expression, *pir* transcription is highly regulated, and it is possible that low-level transcripts exist below the threshold required to mount an immune response, or alternatively, that so few iRBC are targeted by immune responses, that the majority of parasites remain to prolong the infection (Cunningham et al., 2009). Both of these proposals are possible mechanisms for the *pir* family to mediate immune evasion, even if this is not antigenic variation in the strictest sense.

yir expression has been shown to be modulated in immuno-competent, but not immuno deficient mice, suggesting that YIRs are immune targets (Cunningham et al., 2009, Cunningham et al., 2005). In addition, few iRBCs from either laboratory infection, or patient isolates, are recognized by Abs generated against different sub-sets (Fernandez-

Becerra et al., 2005, Cunningham et al., 2005), indicating that there is differential expression at the protein level too. Finally, the fact that there appears to be limited cross-reactivity between Abs generated to YIR sub-family specific peptides (Cunningham et al., 2005), supports the idea that distinct PIR variants could be maintained in the face of an active host immune response.

1.9.5 ii) Immune interactions

As described above, *pir* transcription is modulated in the presence of an intact immune response, suggesting that PIR proteins could be immune targets. Abs are indeed generated to PIRs during *P. yoelii* and *P. vivax* infections (Cunningham et al., 2009, Del Portillo et al., 2001), although the number of *P. vivax* episodes experienced by the patients did not significantly alter the repertoire of VIR that their Abs were able to detect (Fernandez-Becerra et al., 2005, Oliveira et al., 2006). This provides support for the smoke-screen hypothesis of immune evasion, whereby multiple PIRs may be expressed during infection, overwhelming the immune response such that some iRBCs always escape immune selection to proliferate and maintain the infection.

Another potential immune interaction has been postulated for the KIR family members, in that several of these proteins contain regions of identical sequence to those of host proteins (Pain et al., 2008). Notably, the molecules CD99 and AHNAK are included in this phenomenon, where different KIRs contain identity to different regions of the proteins. KIR sequences correspond in total to over half of the extracellular domain of CD99, a protein expressed on all leukocytes, which aids their migration, adhesion and activation (Dworzak et al., 1994). Pain and colleagues suggested that KIR proteins mimic host CD99, which may act to inhibit iRBC recognition by leukocytes or compete for CD99 ligands to reduce T cell activation (Pain et al., 2008). To date similar sequences have not been determined in other PIR members, but it is possible that similar host mimicry may also occur in the other species.

1.9.5 iii) Sequestration

P. falciparum and the rodent parasites *P. chabaudi* and *P. berghei*, are known to sequester in various organs, to differing degrees (Montgomery et al., 2007, Mota et al.,

2000, Franke-Fayard et al., 2005), whilst non lethal clones of *P. yoelii* are thought not to, because all intra-erythrocytic stages can be seen in peripheral blood, also the case for *P. vivax*. A recent study has questioned this dogma however, showing that *P. vivax* iRBCs can cyto-adhere via the receptors ICAM-1 and CSA (Carvalho et al., 2010). In this study, *P. vivax* iRBCs could cyto-adhere with approximately a tenth of the frequency of *P. falciparum* iRBCs, but once bound the interactions were as strong.

The ability of only some iRBCs to sequester would explain the observation of mature stages in peripheral circulation, and also explains their depletion in peripheral circulation found in some human and monkey infections (Rudolf, 1927, Field and Reid, 1956, Fremount and Miller, 1975). Indirect evidence exists that lung pathology may result from sequestered *P. vivax* iRBCs (Anstey and Price, 2007). Both VIR and YIR family members have been predicted to be adhesive (Ansari et al. 2008), and a recent experimental investigation of *P. vivax* iRBC adhesion has implicated the VIR proteins in this process. Anti-sera raised against the VIR sub-families A and E were both able to reduce cyto-adhesion to human lung endothelial cells, *in vitro*, by approximately 40% compared to iRBCs incubated with no anti-sera, and approximately 30% compared to iRBCs incubated with unrelated anti-sera (Carvalho et al., 2010). This is the first evidence that members of the PIR family could mediate cyto-adhesion, although further, more rigorous investigations will be required to elucidate the role of VIR proteins in this phenomenon. However, correlations of PIR with sequestration are complex, as Lovegrove and colleagues found no difference in *bir* transcription when mice were infected with different strains of *P. berghei*, which either did or did not elicit cerebral malaria (Lovegrove et al., 2006). However, 5% of *bir* genes were differentially expressed when CM-resistant and -susceptible strains of mice were compared during *P. berghei* ANKA infection (Lovegrove et al., 2006).

The classification of cerebral malaria in patients, and its representation by rodent models has recently been a subject of debate [for example: (Riley et al., 2010, Stevenson et al., 2010)]. Some data suggest that sequestration of iRBCs is not linked with the pathology of experimental CM (Franke-Fayard et al., 2005), whilst other studies have argued that the extreme inflammation observed in this condition relies upon iRBC sequestration (Lovegrove et al., 2006). What is clear however, is that *bir* transcripts were detected from multiple organs in mice, implying that even if their involvement in CM is not clear, BIR proteins may mediate cytoadhesion in organs such as the lungs, liver and spleen (Lovegrove et al., 2006).

A further hypothesis has been put forward involving splenic sequestration, which would explain the partial lack of mature *P. vivax* iRBC in peripheral circulation, (Del Portillo, 2004, Fernandez-Becerra et al., 2009). During malaria infection, the architecture of the spleen is dramatically altered (Achtman et al., 2003, Cadman et al., 2008). One of the changes is the development of barrier cells (Weiss et al., 1989) between the white and red pulp, to which Del Portillo and colleagues believe that PIR may allow sequestration of *P. vivax* iRBC (Del Portillo et al., 2004).

Clearly there is still a great deal to learn about this large multi-gene family, particularly with respect to their possible host-interactions such as sequestration and immune evasion. These investigations are most accessible in rodent models, following which mechanisms can be confirmed by patient studies.

1.10 Investigations of the *cir* multi-gene family

1.10.1 *P. chabaudi*

P. chabaudi is the only well-characterised *Plasmodium* species that produces chronic infection in laboratory mice - in contrast to both *P. berghei* and *P. yoelii*, which give rise to an acute infection that is either resolved or results in death [reviewed by (Amante and Good, 1997)]. The chronicity of *P. chabaudi* infection produced in laboratory mice is likely to be perpetuated by antigenic variation. Antigenic variation has been demonstrated within this model (McLean et al., 1982b, McLean et al., 1982a, McLean et al., 1986), however the antigens responsible have not been determined.

Besides development of chronic infection, *P. chabaudi* also exhibits several other features observed in human *Plasmodium* infections, including rosetting, sequestration of some mature iRBC stages and adhesion to endothelium via the receptor CD36 (Gilks et al., 1990, Mota et al., 2000, Mackinnon et al., 2002).

Another similarity to human malaria is that mice develop some immunity to re-infection with the same parasite strain after only one infection (Jarra et al., 1986), otherwise known as homologous challenge. However, resistance to heterologous challenge, re-infection using a different parasite strain, did not emerge even after 6 or 7 homologous infections (Jarra et al., 1986). The *P. chabaudi* model thus provides an ideal system in which to dissect the host immune response to *Plasmodium* infection, and this has been extensively investigated, as described in section 1.3.3.

Since many facets of mouse infection with *P. chabaudi* have been elucidated or are experimentally accessible, this system may also be used to investigate which proteins could mediate antigenic variation or sequestration, and which are the targets of protective immune responses.

The *cir* multi-gene family was chosen for the present study. Because the *cir* genes comprise the largest gene family in *P. chabaudi*, their products are prime candidates for antigenic variation, immune evasion and potentially sequestration, yet very little is known about their expression, function and role in stimulating or evading host immunity.

1.10.2 Objectives of this study

To appreciate the role of CIR proteins during *P. chabaudi* infection several components had to be considered, including the gene-family composition, their expression at the gene and protein level, and their interactions with the host immune response. A wide range of approaches was required to address such different aspects, reflected in the five results chapters presented here:

Firstly, *cir* genes were identified and annotated onto the *P. chabaudi* AS genome sequence, in collaboration with the Sanger Institute. The amino acid sequences were aligned to identify groups of similar sequences, which may share functional roles.

Analysis of *cir* transcription was undertaken, initially by quantitative RT-PCR using primers designed to amplify members of the individual groups; and subsequently methods of purifying *P. chabaudi* AS infected blood were investigated to provide a clean template for full transcriptome sequencing by Solexa / Illumina technology.

Highly transcribed *cir* genes were selected for the design of synthetic gene constructs, used for the expression of recombinant CIR in the methyltropic yeast *Pichia pastoris*.

Polyclonal anti-sera were generated in rabbits against two recombinant CIR proteins and against three CIR peptides. One of these peptides was a conserved sequence found in the majority of CIRs, and anti-sera to this peptide was used to show the localization of CIR proteins within and on the surface of infected erythrocytes.

Finally, the CIR peptides and recombinant proteins were used to determine whether mice make antibodies that recognize CIR proteins during infection, and whether the induction of high anti-CIR antibody titres by immunization might play a role during a subsequent *P. chabaudi* AS infection.

Chapter 2: Characterization of the *cir* gene family

2.1 Introduction

The sequencing of *Plasmodium* genomes has made it possible to apply bio-informatic methods towards characterizing novel genes. Several approaches have been used in efforts to identify and understand the *pir* genes, described below.

The identification of thirty-two novel genes in a *P. vivax* yeast artificial chromosome library clone marked the discovery of the *vir* multi-gene family (Del Portillo et al., 2001). Shortly afterwards, the *cir* genes were identified in *P. chabaudi*, by a genome survey sequencing project, and initially called *fam3* (Janssen et al., 2001, Fischer et al., 2003, Janssen et al., 2002). Orthologues were also detected in the available genome sequencing data for the rodent parasites *P. berghei* and *P. yoelii,* whereupon the whole 'super-family' was called *pir* (Janssen et al., 2002, Janssen et al., 2004, Carlton et al., 2002).

These early investigations were unable to use complete genome sequences for identification of the full *pir* repertoires, and instead estimated the total copy number according to the distribution of *pir* genes within the investigated DNA and the fact that all chromosomes were positive by southern blots using *pir* specific probes. Initially, 600- 1000 *vir* genes (Del Portillo et al., 2001) and 600 *cir* genes (Janssen et al., 2002) were predicted, figures which were revised closer to the true copy numbers as genome data accumulated: 346 *vir* genes, 38 *kir* genes, 160-200 *bir* and *cir* genes, and 838 *yir* genes were predicted (Carlton et al., 2002, Janssen et al., 2004, Pain et al., 2008).

In addition, chromosomal location made annotation of *pir* genes particularly difficult. The majority of *pir* genes reside in the sub-telomeres (Fischer et al., 2003, Carlton et al., 2008, Hall et al., 2005, Carlton et al., 2002), which are prone to repetitive sequence (Hall and Gardner, 2004), or associated with interstitial repeat sequences making contig assembly in these regions fraught with error (Hall and Gardner, 2004). For *Plasmodium* species whose genomes are still incomplete, such as *P. berghei* (according to http://www.genedb.org/Homepage/ Pberghei 4/4/2011), the *pir* annotation is likely to change as genome-finishing re-assembles contigs and accurately defines the sub-telomeric and repetitive sequences.

As genome-sequencing data expands, the difficulties of *pir* chromosomal location may

be resolved, such as for the *vir* and *kir* genes (Pain et al., 2008). However, automated gene prediction software remains unable to accurately detect all *pir* genes, meaning that, to date, the majority of *pir*s have been manually annotated [for example, (Carlton et al., 2008)].

Whilst determination of the full genome sequence has allowed description of the *vir* and *kir* families in detail (Pain et al., 2008), and smaller scale investigations have indicated the conserved features of PIR superfamily members (Del Portillo et al., 2001, Janssen et al., 2002, Janssen et al., 2004, Merino et al., 2006, Fonager et al., 2007), less is known about the *pir* repertoire in rodent malarias. The *P. chabaudi* AS genome had been sequenced to 8-fold coverage at the beginning of this study, making it possible to update the *cir* repertoire from its original annotation (Janssen et al., 2002).

The objectives of the work described in this chapter were to:

i) Identify and annotate the full *cir* repertoire in the *P. chabaudi* AS genome.

ii) Characterize the features of CIR amino acid sequences that may provide indications to the possible function(s) of these proteins

2.2 Methods

2.2.1 Annotation of *cir* genes

Artemis release 11 (Rutherford et al., 2000, Carver et al., 2008) was used to annotate *cir* genes onto the 8x sequence coverage of the *P. chabaudi* genome. This process is summarized in Figure 13.

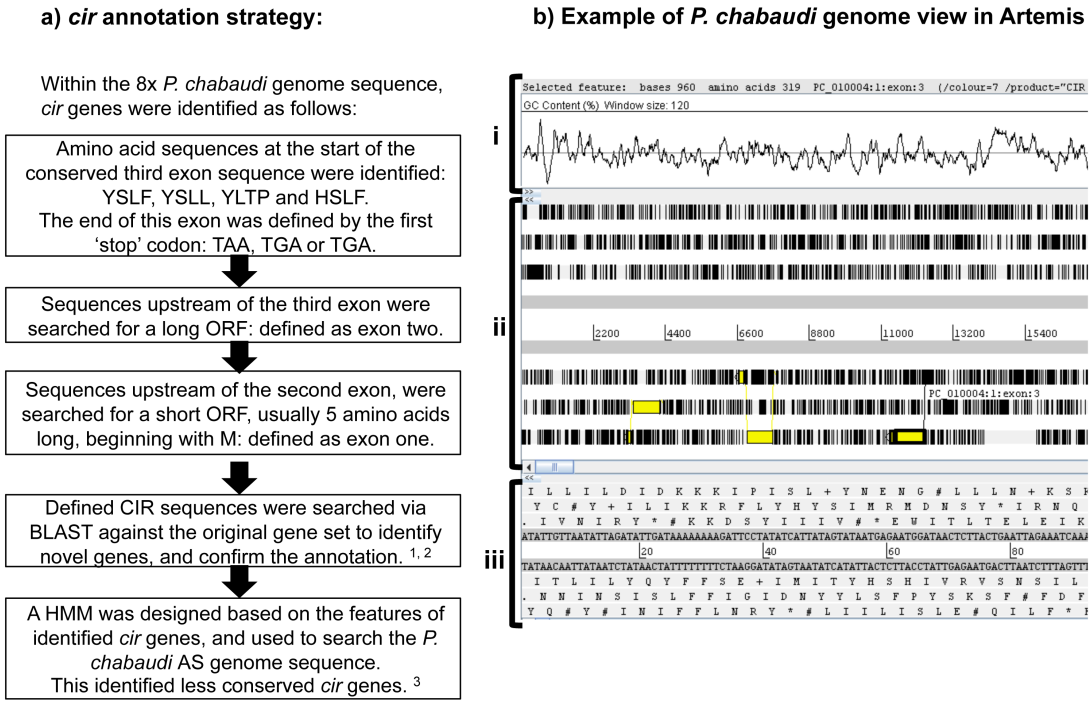


Figure 13: Strategy used for annotation of *cir* genes in the *P. chabaudi* AS genome.

- **a)** Flow chart demonstrating the steps taken to identify and annotate *cir* genes.
- **b)** Annotation was carried out using the genome browser Artemis (Rutherford et al., 2000, Carver et al., 2008). Three windows show important features of the selected DNA assembly:
  - **i.** GC content is shown graphically, fluctuating around the average content for each DNA section under analysis.
  - **ii.** All six possible open reading frames are shown in yellow, with stop codons represented by black vertical bars. Different exons often existed in different reading frames, as shown in the three example *cir* genes included in this DNA sequence.
  - **iii.** All six possible reading frames are shown, as in window **ii)**, but further magnified in this view, to show the individual amino acid sequences.

Three stages of *cir* gene annotation were performed, indicated in Figure 13 by [1], [2], and [3]. Open reading frames (ORFs) were identified that contained conserved features of *cir* genes from the previous (3x) genome assembly and common *Plasmodium* splice sites (Appendix 2.1). Available *P. chabaudi* AS transcriptome data was used to verify the annotated ORFs. These were then searched for similarity to the *pir* super-family genes via basic local alignment tool (BLAST). Genes that had high identity to other *pir* genes were considered to be *cir*s.

110 *cir* genes were originally identified, which increased to 128 *cir*s upon genome finishing where contigs were re-assembled. This enabled the re-evaluation of *cir* genes which had either previously been found as partial genes at the ends of contigs, or had been mis-identified within poorly assembled regions of the genome. A further 71 *cir* genes were identified using a hidden Markov model [HMM, reviewed by (Eddy, 1996)], created using the established repertoire of 128 *cir* genes, according to the methods used for Pfam classification (Sonnhammer et al., 1998, Finn et al., 2010).

2.2.2 Detection of conserved motifs

CIR amino acid sequences were uploaded to http://meme.sdsc.edu/meme/ to allow Multiple Expectation Maximization for Motif Elicitation analysis [MEME, (Bailey and Elkan, 1994)], and the program was set to identify up to 20 of the most conserved motifs and return the analysis as text. The average motif locations were identified and plotted onto each gene. WebLogos were generated using by the MEME program, using an adaptation of the WebLogo software [(Crooks et al., 2004), hosted at http://weblogo.berkeley.edu/].

2.2.3 Analysis of sequence similarity

Three stages of annotation were performed, outlined in Figure 13; the number of identified *cir* genes increased after each stage of annotation. Amino acid sequences of 107, 117 and 183 CIRs were aligned using the Multiple Sequence Comparison by Log-Expectation algorithm [Muscle, (Edgar, 2004)]. These alignments are attached in Appendices 2.3, 2.6 and 2.9. (In addition, alignments of PIR sequences are attached in Appendices 2.12 and 2.14). Certain sequences aligned poorly with the other CIRs, for reasons including probable incomplete annotation and extreme divergence (called 'CIR-

like' sequences). These sequences were excluded from each alignment. Alignments were then refined by eye to ensure that all the sequence blocks identified aligned well together. Regions containing large insertions were deleted as these cannot be compared to the other sequences and are therefore uninformative.

Several tree-building methods were applied to the first alignments of 107 and 117 CIRs, to ensure that the clades seen were valid. Phylogenetic trees were produced with the MEGA program (Kumar et al., 2004) using: un-weighted pair group method with arithmetic mean (UPGMA), minimum evolution, neighbour joining and maximum parsimony. The latter three trees used an evolutionary model for amino acid sequence, the Poisson correction (Zuckerkandl and Pauling, 1965). For these methods a bootstrap of 500 was carried out. This calculates the proportion of times a particular branch was included in a tree, so a bootstrap of 50 means 50% of trees contained that branch (Felsenstein, 1985).

In addition, a tree was constructed using the maximum likelihood method from the PhyML server (Guindon et al., 2005, Guindon and Gascuel, 2003). Here, the evolutionary model applied was Le Gascuel [LG, (Gascuel, 1997)] and the branch support was calculated by approximate likelihood ratio test [aLRT, (Anisimova and Gascuel, 2006)]. This method is a non-parametric branch support based on a Shimodaira-Hasegawa-like procedure [SH-like, (Shimodaira and Hasegawa, 1999)], which is less computationally intensive than bootstrapping, but not equivalent. SH-like aLRT assesses whether the branch being studied is more statistically likely than collapsing a branch, leaving the rest of the tree topology identical. The initial tree was created by BIONJ (Gascuel, 1997), and tree improvement was carried out by nearest neighbor interchanges (NNI) using the maximum likelihood method.

Clades that were identified by all of these analyses, with high branch support values, contained highly CIR similar sequences; consequently these clades were called subfamilies A - F. The maximum likelihood algorithm implemented by PhyML (Guindon and Gascuel, 2003, Guindon et al., 2005), was seen to provide more consistent results than the other tree-building methodologies, therefore only this method was applied to the alignments of 117 and 183 CIRs. Maximum likelihood trees are attached in Appendices 2.4, 2.7 and 2.10.

Since trees only represent a bifurcating lineage and imply phylogenetic relationships, reticulate networks were also created within the program Splitstree 4.0 (Bryant and

Moulton, 2004). All networks used the algorithms NeighborNet (Bryant and Moulton, 2004) for calculation of distances and Equal angle (Gambette and Huson, 2008) for calculation of splits, as these make no assumptions about the evolutionary history of sequences, and are therefore suitable for creation of character display networks (Bryant and Moulton, 2004, Morrison, 2005). 1000 bootstrap replicates were generated.

NeighborNet (Bryant and Moulton, 2004, Morrison, 2005) is derived from the Neighbour joining tree-building method, and uses agglomeration to produce progressively larger overlapping clusters of taxa. This algorithm is particularly suited to complex data sets, such as the CIRs, as fewer false negatives are produced than other commonly used methods of network creation, such as split decomposition (Bryant and Moulton, 2004, Morrison, 2005). Networks created for each CIR alignment are attached in Appendices 2.5, 2.8 and 2.11, whilst networks of PIR sequences are attached in Appendices 2.13 and 2.15.

2.2.4 Evidence of recombination between *cir* genes

Phylogenetic incompatibilities within the alignment of 183 CIRs, and smaller alignments containing members of each identified sub-family, were analysed using the pairwise homoplasy index (PHI) in Splitstree v4.0 (Huson, 1998).

Phylogenetic profiling was used to detect phylogenetic inconsistencies between four chosen *cir* DNA sequences, using a hidden Markov model method within the TOPALi platform (Milne et al., 2009, Milne et al., 2004). The probability of generating each of the three possible tree topologies for the four sequences was modelled in a given 100 nucleotide window. Possible recombination breakpoints were identified where the most probable topology altered at different positions along the alignment.

2.2.5 Function shift analysis

The two major CIR sub-families (U and A-F) were compared by identification of residues in the alignment that were differently conserved between the two groups. The 'FunShift' methodology was developed using enzymes known to have different specificities (Abhiman and Sonnhammer, 2005), and is thus designed to predict whether groups of proteins may perform different functions.

The alignment of 183 CIRs was split into two files containing only CIR sub-family A-F or CIR sub-family U members. The two alignments were then used to calculate the following parameters, described below. Positions that contained only gaps in a subfamily were not counted.

- Rate-Shifting Sites (RSS) were defined as positions conserved in one sub-family but variable in the other, and were identified using the likelihood ratio test (LRT) program [(Knudsen and Miyamoto, 2001), hosted at http://www.daimi.au.dk/~compbio/rateshift/]. The U-values generated by this program indicate the likelihood of rate change for each position in the alignment between the two sub-families. U-values above 4.0 were considered significant at the 5% significance level, as previously described (Knudsen and Miyamoto, 2001).
- Conservation-Shifting Sites (CSS) were defined as positions that were conserved in both groups, but containing different residues in each. CSS were detected using the method developed by Abhiman, Sonnhammer and colleagues [(Abhiman and Sonnhammer, 2005), hosted for proteins defined by PFAM domains at http://funshift.sbc.su.se/]. This calculates a Z-score based on the normalized cumulative relative entropy at each position of the alignment, between the two sub-families. Z-scores exceeding 0.5 per alignment position were considered significant (Abhiman and Sonnhammer, 2005).

2.3 Results

2.3.1 Identification of *cir* genes

Completion of the *P. chabaudi* AS genome sequencing to eight fold redundancy meant it was possible to revisit initial *cir* annotation (Janssen et al., 2002, Janssen et al., 2004). Conserved features from initially identified *cir* genes, such as the relative exon lengths, splice sites and amino acid sequences were used to identify which ORFs in the assembled contigs were likely to be *cir* genes (This information is included in Appendix 2.1). Three stages of annotation were performed, outlined in Figure 13; the number of identified *cir* genes increased after each stage of annotation.

At first, 110 *cir* genes were identified, and 56% of these could be mapped to a specific chromosome. BLAST analysis found that 44% of newly annotated *cir*s were identical to *cir*s annotated on the previous genome assembly, which had a three-fold depth of sequence coverage. Some previously identified *cir* genes had high similarity to more than one newly annotated *cir*, in particular Pc_200009 (whose current ID is PCHAS_000100) matched 8 *cir*s with greater than 73% identity. The expectation (E-) value of this tBLASTn match is 7e-108, the low value indicates that this is a significant alignment and unlikely to have occurred by chance.

As the final *P. chabaudi* AS contigs were assembled accurately, a further 18 *cir* genes were detected. Finally, a hidden markov model (HMM) was designed on the basis of the identified *cir* repertoire (Sonnhammer et al., 1998, Finn et al., 2010), which was used to detect more divergent *cir* genes, bringing the total of identified *cir*s to 198. These data are summarized in Table 5.

Table 5: The status of the *cir* repertoire at each stage of annotation.

Annotation stage:	1	2	3
Number of annotated *cir* genes	110	128	198
Long *cir*s	Pc_040001 Pc_990013 Pc_990028 Pc_990042	PCAS_001000 PCAS_000210 PCAS_000140 PCAS_040020	PCHAS_040020 PCHAS_000130 PCHAS_000400
Pseudogene	Pc_070004	0	PCHAS_070090
Number of *cir*s with no predicted TM domain	0	0	8
Number of *cir*s with 1 predicted TM domain	110	123	143
Number of *cir*s with 2 predicted TM domains	0	5	36
Number of *cir*s with >2 predicted TM domains	0	0	1 (PCHAS_011490)

The species identifiers: Pc_, PCAS_, PCHAS_ and gene names were altered at each stage of annotation.

Only three partial *cir* genes were identified in each stage of annotation, Table 5. These may result from poor assembly of contigs that could not be assigned to a specific chromosome. Long *cir* genes were also identified, which contained an extended first exon. The majority of *cir* genes contained one predicted TM domain, even after more divergent *cir*s were identified. Upon addition of more divergent *cir* genes to the repertoire, some amino acid sequences aligned poorly with the majority of CIRs, contained a long third exon or had been indentified as CIR-like sequences, containing some but not all expected features of *cir* family members. These CIR sequences were excluded from further analysis.

2.3.2 CIR sequence similarity

One method for further characterization of the *cir* repertoire is to investigate sequence similarity. To this end, the amino acid sequences detected at each stage of *cir* annotation were aligned using Muscle (Edgar, 2004), and refined by eye. The sequence similarity of each alignment was determined using Plotcon (http://emboss.bioinformatics.nl/cgi-bin/emboss/plotcon), Figure 14. Clearly, all three alignments were similar, with more conserved sequences found between amino acids 250 and 600, and at the end of the alignment. The length of the alignments differed predominantly with respect to the region encoded by the end of exon two, which increased as more divergent CIRs were included.

To determine the relationships between CIR sequences, both phylogenetic and non-assumptive methods were employed. Initially, phylogenetic analysis was carried out using several methods from the alignment of 107 CIRs. The ability of different phylogenetic methods to identify clades of similar CIR sequences is compared in Figure 16. Three distance-based phylogenetic methods were used: un-weighted pair group method with arithmetic mean (UPGMA), a, minimum evolution, b, and neighbour joining, c (Figure 15). These are simple methods that rank sequences according to the number of pairwise differences at each point, and then use this distance matrix to calculate a phylogeny. UPGMA is the simplest, followed by minimum evolution, which estimates the length of each branch in the tree, and chooses the ‘best’ tree as the one with the lowest sum of branch lengths. Neighbour joining is slightly more complex as an evolutionary model is applied to calculate the tree.

Two cladistic methods were also used: maximum parsimony, d, and maximum likelihood, e, (Figure 15). These methods create an initial tree directly from the multiple sequence alignment so no information is lost. Many trees are created by these methods, and the best one is chosen according to an evolutionary model. Maximum parsimony chooses the best tree as the one that uses the fewest evolutionary changes. Maximum likelihood calculates the probability of all possible tree topologies for each sequence unit, and chooses the most likely tree with the highest overall probability at each residue. All of the phylogenetic methodologies gave rise to similar tree morphologies, where 76% of CIRs clustered to form 6 clades, labelled A-F, Figure 15. The branches were coloured according to sub-family. Clades were only deemed to be sub-families if the bootstrap or aLRT branch support value was greater than 60%, in the majority of trees. This was the case for sub-families A, B, D, E and U, and the C_1, C_2, F_1 and F_2 clades. Sub-family C, and the clades F_1 and F_2 were all strongly supported in the maximum likelihood tree, e) Figure 15, with aLRT values of 95 or greater.

In addition, networks were created in the program Splitstree 4 (Huson and Bryant, 2006) using the NeighborNet and Equal Angle algorithms (Bryant and Moulton, 2004, Gambette and Huson, 2008), for all alignments of CIR sequences. Figure 16 shows the trees and networks created from alignments of 117 and 183 CIRs. Unlike the phylogenetic analyses, the network methodology used did not assume linear evolution, instead allowing the complex inter-relationships between sequences to be visualized, which may reflect their evolutionary history or simply display character conflicts within the data [reviewed by (Morrison, 2005)]. The maximum likelihood trees and NeighborNet networks produced clades with consistent confidence values, despite being based on different criteria. The figures with confidence values (either calculated by bootstrapping, for networks, or aLRT, for maximum likelihood trees) are attached as appendices.

Although the six sub-groups A-F were consistently identified (Figure 15 and Figure 16), their significance dropped as more divergent CIR sequences were added to the alignment. From the largest alignment of 183 CIR sequences, the sub-groups A-F comprised 44.3% of all analyzed CIRs. In contrast, the un-clustered CIR sequences, highlighted in black, expanded greatly upon identification of further CIR sequences by HMM analysis. This was expected given the divergent nature of the un-clustered CIRs, however clades of similar sequences were present within this group, Figure 16. Due to time constraints these clades were not investigated further, and the 55.7% CIRs found in

the un-clustered clade were termed sub-family U. The members of each identified sub-family are shown in Table 6.

Table 6: Genes present in the CIR sub-families: A, B, C, D, E, F and U.

Sub-families									
A	**B**	**C**	**D**	**E**	**F**	**U**			
140030	000400	114750	120070	011510	140090	000040	083760	042020	040060
114740	040020	000140	001130	104260	000730	000390	060050	060070	000430
000030	120040	000070	100040	140020	030020	001100	030090	030190	104200
000090	000130	001040	000560	070020	000360	000500	060060	060090	114720
000270	001110	030080	000300	000720	011530	030120	130030	041980	073130
000120	000110	060020	073180	070050	073160	114640	042030	130080	030070
000280	120050	000020	070130	030060	070160	030180	060140	041970	137110
000490	104250	000290	040030	000750	000580	130070	040110	060110	011490
010020	000320	001090		000680	010030	146850	120030	041990	070060
140130	000410			000100	110020	011500	000660	030140	070040
000570				000310	050060	070100	040080	130100	011480
114730				000420	000350	030110	060160	070170	000470
140040				040040	050020	130050	146790	100060	050040
100030				120060	073200	040050	000260	130060	000770
001120				104230	120020	001050	083720	146860	130220
000740				030040	000060	050070	090010	060130	114700
011520				000340	130020	140070	114600	041950	137030
070030					000150	000180	042070	130120	000170
073190						001060	010040	126770	130170
						011450	073150	130280	070070
						030210	140140	011330	000220
						146870	110030		
x 19	x 10	x 9	x 8	x 17	x 18	x 102			

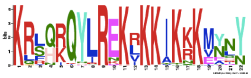
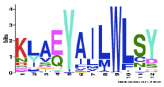
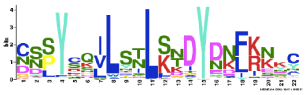
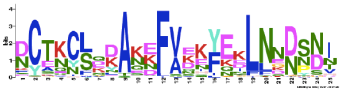
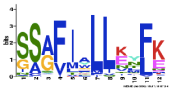
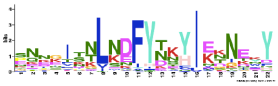
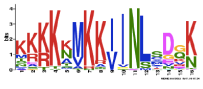
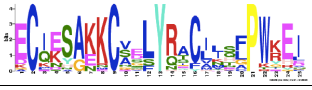
Sub-families were identified from the alignment of 183 CIRs. The number of *cir*s present in each sub-family are indicated underneath the gene names. For clarity, the species identifier PCHAS_ was excluded for each gene.

Given that the later analyses of the CIR repertoire appeared to show two much larger clusters of sequences than previously identified, each comprising almost half of the CIRs, these may be considered to be 'major' sub-families. For simplicity, the major CIR sub-families will henceforth be referred to as sub-family U (comprising the more divergent CIRs) and sub-family A-F (including all six of the minor sub-groups that were initially identified), as shown in Figure 16.

2.3.3 Identification of conserved amino acid motifs

The identification of conserved amino acid sequences may provide indications for possible protein function(s). 14 conserved motifs in the CIR amino acid sequences were identified by Multiple EM for Motif Elicitation analysis [MEME, (Bailey and Elkan, 1994)]. These are represented as WebLogos (Crooks et al., 2004) in Table 7, where the height of each letter indicates the proportion of CIR sequences containing that residue.

Table 7: Conserved amino acid motifs found within CIR sequences.

Number	Motif	Frequency	E-value
1		72.73%	2.1e-1491
2		33.84%	3.9e-939
3		72.22%	4.2e-923
4		74.24%	3.7e-1228
5		72.22%	2.2e-1262
6		39.80%	2.6e-628
7		41.41%	4.0e-800
8		15.66%	3.0e-426
9		49.49%	6.8e-391
10		59.09%	4.6e-567
11		45.45%	6.5e-462
12		29.80%	4.6e-337
13		12.63%	2.6e-326
14		14.14%	7.5e-321

Amino acid motofs identified by MEME analysis (Bailey and Elkan, 1994). Hydrophobic residues are shown in blue, polar, non-charged residues in green, acidic residues in pink and positively charged residues in red. The frequency is the percentage of motif occurrences within the 198 analyzed CIR sequences. The E- values refer to the significance of each motif found within the CIR sequences.

Motifs 1 and 3 were the most conserved, being present in 72% of CIRs, and containing particular residue conservation. Motif 1 contained the almost unanimous YK residues, corresponding to the start of the third *cir* exon and part of the predicted CIR trans-membrane domain. Motif 3 was found within the second *cir* exon and contained the highly conserved sequence YAILWLSY. Motif 10, present in 59% of CIRs, contained some degeneracy, but a clear conserved methionine indicated the N-terminus of all CIR proteins. All CIRs possessing motif 10 also contained a cysteine six residues into the sequence. Conserved cysteine residues were also located within motifs 5, 7 and 14.

CIR motifs were arranged according to their phylogenetic sub-group to determine whether any motifs were sub-family specific, shown in Figure 17. Motif 10 was found at the N-terminus of all CIR proteins. Motifs 9, 3 11, 5, 4 and 1 were also found in members of all sub-families, indicating that these amino acid sequences may be essential for all CIR proteins. The remaining motifs appeared to have sub-family specific distributions, with motif 2 specifically, and motifs 6 and 7 predominantly found within sub-families A-F. By contrast, motifs 8, 12, 13 and 14 were specific to CIR sub-family U. Motif 8 lies within the predicted TM domain, which indicates that CIR sub-family U proteins may have different functions than the other CIRs, such as perhaps spaning a different membrane.

2.3.3 Diversity in the *cir* family generated by recombination

The detection of similar CIR sequences, with conserved and sub-family specific amino acid motifs indicates that genetic exchange can occur between members of this multi-gene family, most likely by recombination.

Different types of recombination exist, several of which may affect *pir* genes. Homologous recombination occurs between similar sequences, typically alleles of the same gene, which could extend further within the *pir* repertoire to include any *pir* containing similar sequence. This may therefore occur during mitosis (within the mammalian host) as well as meiosis (within the mosquito), since the haploid genome present in all mammalian stages of the *Plasmodium* lifecycle does not necessarily preclude recombination between similar *pir* genes. During homologous recombination, genetic material may be exchanged in a reciprocal manner or one chromosome may donate genetic material without itself being changed. This phenomenon is called gene

conversion, and is likely to affect all multi-gene families. For simplicity, and since it is difficult to distinguish between these events bio-informatically, both reciprocal recombination and gene conversion will be referred to as 'recombination'.

Visualization of CIR sequence relationships using network methodology indicated that recombination may have occurred within the CIR repertoire, as many box-like structures (or reticulations) were present, Figure 16. Generally, tree-like structures within the network indicate linear evolution; whilst reticulations may represent recombination (Morrison, 2005). However, the interpretation of networks is more complex than interpretation of dichotomous trees. In fact, reticulations may represent one of three possibilities (Morrison, 2005):

- Uncertainty or ambiguity in the data (for example: mistaken sequence homology, mis-alignment, or a poor fit of the data to the model used)
- Analogy events in evolution (for example: convergent evolution giving rise to similar residues in sequences without shared origins)
- True genetic exchange events in evolutionary history (for example: recombination, lateral transfer or hybridization)

The second possibility may be ruled out, as all *cir* genes studied here were present within the same genome. Reticulations within the CIR networks were thus most likely to represent either poor alignment of the amino acid sequences or recombination events.

To determine whether recombination had occurred within the *cir* repertoire of *P. chabaudi* AS, a closer examination was carried out using a variety of bio-informatic methods. An example is illustrated in Figure 18. Sub-family E was chosen since this appeared to be the most conserved of CIR clades, with a longer branch separating this from the other families within this major subgrouping than families A, B, C, D or F (Figure 16b). The members of this sub-family were aligned, and a network generated, as described previously, Figure 18a. The sequence similarity is shown in Figure 18b.

There appeared to be two conserved clades within subfamily E, of which only one retained a significant probability of recombination (P = 0.0) using the PHI test within Splitstree (Huson, 1998) when the other clade was removed. This clade was present in the left hand region of Figure 18a, from which four *cir* DNA sequences were chosen: PCHAS_000100, PCHAS _000310, PCHAS _120060 and PCHAS _040040. These were aligned as previously described for CIR sequences, Figure 18c.

The program 'tree TOPology-related analysis of ALignments interface' [TOPALi; (Milne et al., 2009, Milne et al., 2004)] was used to produce a phylogenetic profile along the length of these sequences, iteratively removing sequence information from either the C-terminus or N-terminus end. By this approach, it may be possible to find areas of recombination between the *cir* sequences, as the tree topology would change when a recombination break-point was exposed. Phylogenetic inconsistencies were observed, indicating that different regions of the genes displayed different relationships with each other, Figure 18c. PCHAS_000100 is shown as an example, Figure 18d, containing similarity to different *cir*s along its length. Therefore it is likely that the PCHAS_000100 gene has undergone several recombination events.

In fact, no part of the *cir* repertoire appeared exempt from recombination. All sequences investigated produced significant results from the PHI test and indicated possible break-points by TOPALi analysis, they do indicate that recombination has played a strong role in shaping the *cir* repertoire.

2.3.4 The context of CIR within the PIR super-family

Since the PIR super-family is present in many *Plasmodium* species, it is likely that these proteins play similar roles, although functional differences may have emerged after speciation. PIR sequences that are conserved between different species may be subject to functional constraint, thus further investigation of such proteins may indicate the conserved role of PIR proteins.

Firstly, a selection of PIR representing many different sub-groups was aligned as before, and used for creation of a NeighborNet network to illustrate the overall sequence diversity within the PIR super-family, Figure 19a (Cunningham et al., 2010). Here, PIR proteins were clearly most similar to PIR from parasites which had undergone the least phylogenetic separation, with VIR and KIR sequences being fairly divergent, and rodent PIRs clustering together, especially the BIR and YIR sequences.

Some CIR sequences appeared to display similarities with BIR and YIR family members, so a larger proportion of the rodent PIR repertoire was analysed according to the same methodology, Figure 19b. Notably, these alignments were of lower quality than those created using only the CIR sequences, so only cautious interpretation may be attempted from this figure. The sub-families represented in these alignments are listed

in Table 8. CIR sequences preferentially clustered with themselves forming two large clades, as seen previously (Figure 16). CIR sub-family U was particularly distinct from the other rodent PIR. Other clades that could be easily distinguished corresponded to 3 of the 5 previously identified YIR sub-groups: G2, G4a, and G5 (Fonager et al., 2007). Each YIR sub-group contained a few similar BIR sequences, but no CIRs. The rest of YIR and BIR sequences clustered together, and contained less distinctive sequences.

Table 8: PIR sequences analyzed in Figure 19, and their respective sub-families.

	PIR sequences		Rodent PIR sequences	
	Number of sequences	**Sub-families of sequences**	**Number of sequences**	**Sub-families of sequences**
VIR	29	B, C, E, I, J, K & nc	-	-
KIR	16	*	-	-
BIR	31	*	83	*
YIR	32	G1, G2, G3a, G5 & nc	274	G1, G2, G3, G4, G5 & nc
CIR	28	B, C, D, E, F & U	143	A, B, C, D, E, F & U
Total	136		500	

Where sequences that were not clustered in previous analyses are labelled nc, and PIR repertoires for which sub-families have not been identified are labelled *. Sub-families that were not analysed are labelled -.

2.3.5 Similarities between CIRs and RIFINs

Whilst the CIR repertoire displayed some similarity with other PIR sequences (Figure 19), CIR sub-family U was distinct from the other rodent PIRs. The identification of two major CIR sub-families (Figure 16) was similar to the organization of the RIFIN repertoire in *P. falciparum* (Petter et al., 2007, Joannin et al., 2008). Accordingly, more comparisons of the CIR and RIFIN repertoires were undertaken.

The RIFIN sub-family A is defined by the presence of a 25 amino acid sequence, which is absent from RIFIN-B types (Petter et al., 2007, Joannin et al., 2008). A relatively conserved insertion was detected only in CIR sub-family U members between position 220 and 316 of the alignment of 183 CIRs. This region of the alignment is shown in Figure 20a, comparing representative CIR sub-family U and sub-family A-F sequences. The same region is also displayed as a weblogo (Crooks et al., 2004), created from all members of CIR sub-family U and sub-families A-F (Figure 20b). Figure 20c shows an enlargement of a section of the CIR sub-family U specific sequence, which displayed some similarities with the A-type RIFIN sequence, included in Figure 20d (Petter et al.,

2007, Joannin et al., 2008). Notably, both sequences included two conserved cysteine residues. Approximately a third of the residues in each insertion sequence were hydrophobic and very few basic residues were present.

The two RIFIN sub-families have been shown to display different sub-cellular localizations (Petter et al., 2007). A- and B-type RIFINs are thus predicted to have undergone functional divergence, which has been supported by bio-informatic analyses of the whole RIFIN repertoire in the *P. falciparum* clones 3D7, DD2 and HB3 (Abhiman and Sonnhammer, 2005). To investigate whether the members of each major CIR sub-family could also have functionally diverged, a similar analysis was performed to that described by Joannin and colleagues (Joannin et al., 2008). The method used for function shift prediction, 'FunShift', has been developed using families of enzymes which were known to have different substrates (Abhiman and Sonnhammer, 2005). This method requires identification of rate shifting sites (RSS) and conservation shifting sites (CSS). These parameters correspond to residues that are either differentially conserved between the two sub-families (RSS), or positions that are equally conserved, but with different amino acids in each sub-family (CSS).

The alignment of 183 CIRs was split into the major sub-families U and A-F; RSS and CSS were compared between the sub-families (significant sites are attached in Appendices 2.16 and 2.17). 77 RSS (15.4% of all positions) and 158 CSS (31.8% of all positions) were identified along the alignment, strongly suggesting that functional divergence between the CIR sub-families may have occurred according to the criteria generated with protein families of known function (Abhiman and Sonnhammer, 2005).

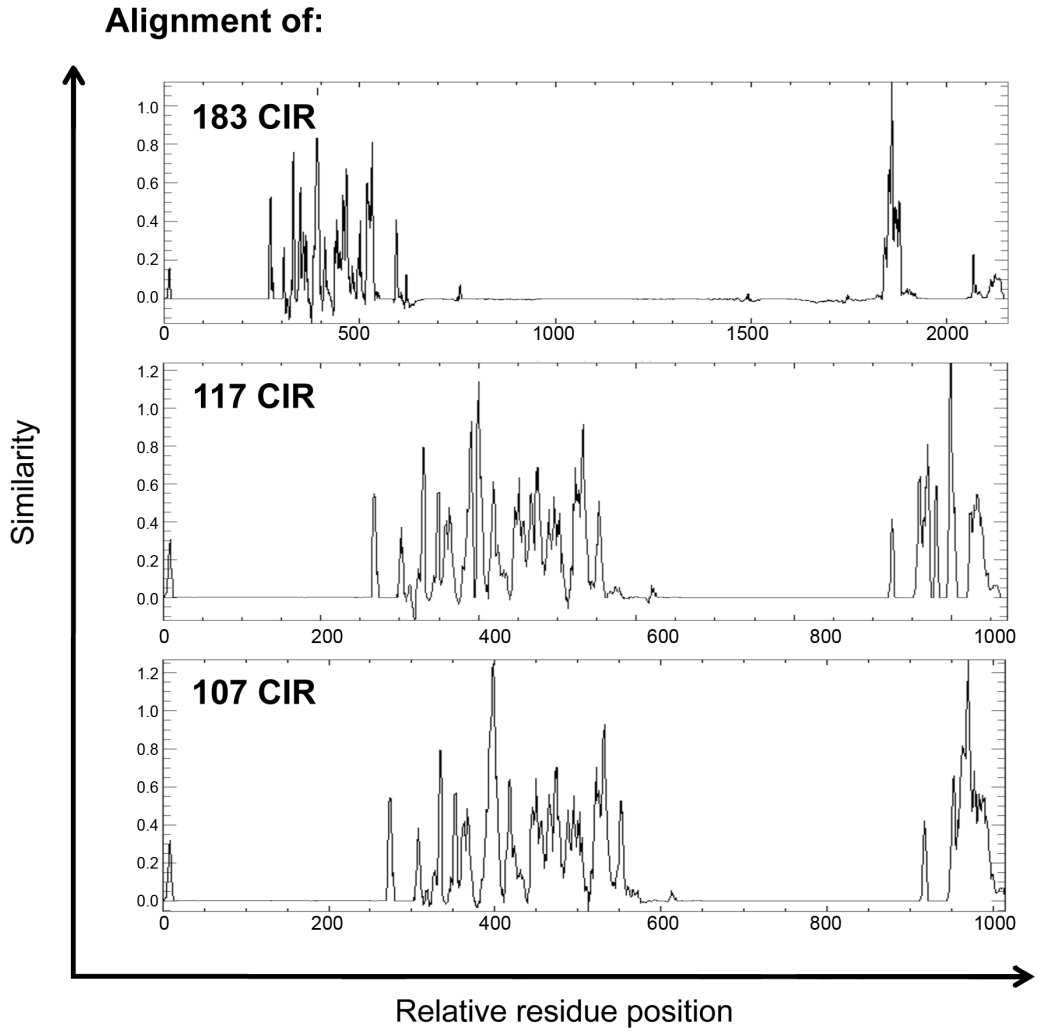
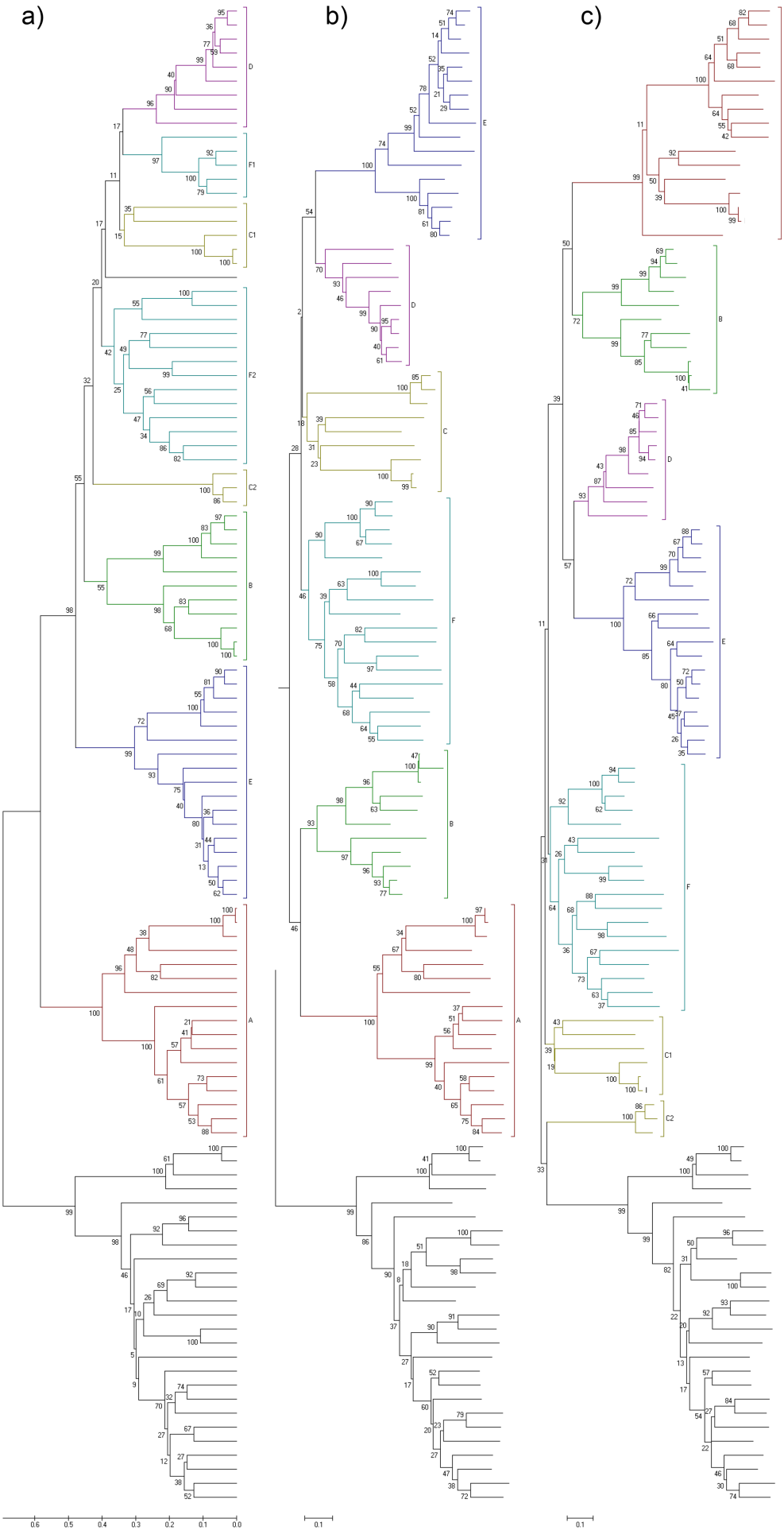


Figure 14: Comparison of alignments using CIR amino acid sequences

At each of the three stages of *cir* gene annotation, the CIR amino acid sequences were aligned using Muscle (Edgar, 2004). Sequences that aligned poorly were removed from each alignment, leaving 107, 117 and 183 CIRs in the alignments at the first, second and third stages of annotation, respectively.

Sequence similarity was plotted for the three alignments of CIR amino acid sequences using Plotcon (http://emboss.bioinformatics.nl/cgi-bin/emboss/plotcon).

As the earliest identified CIRs were more similar, the alignment of 107 sequences was shorter than the subsequent alignments. Gaps present in the alignment are indicated by low sequence similarity.


a)
b)
c)
A
B
C
C1
C2
D
E
F
F1
F2
0.6 0.5 0.4 0.3 0.2 0.1 0.0
0.1
0.1

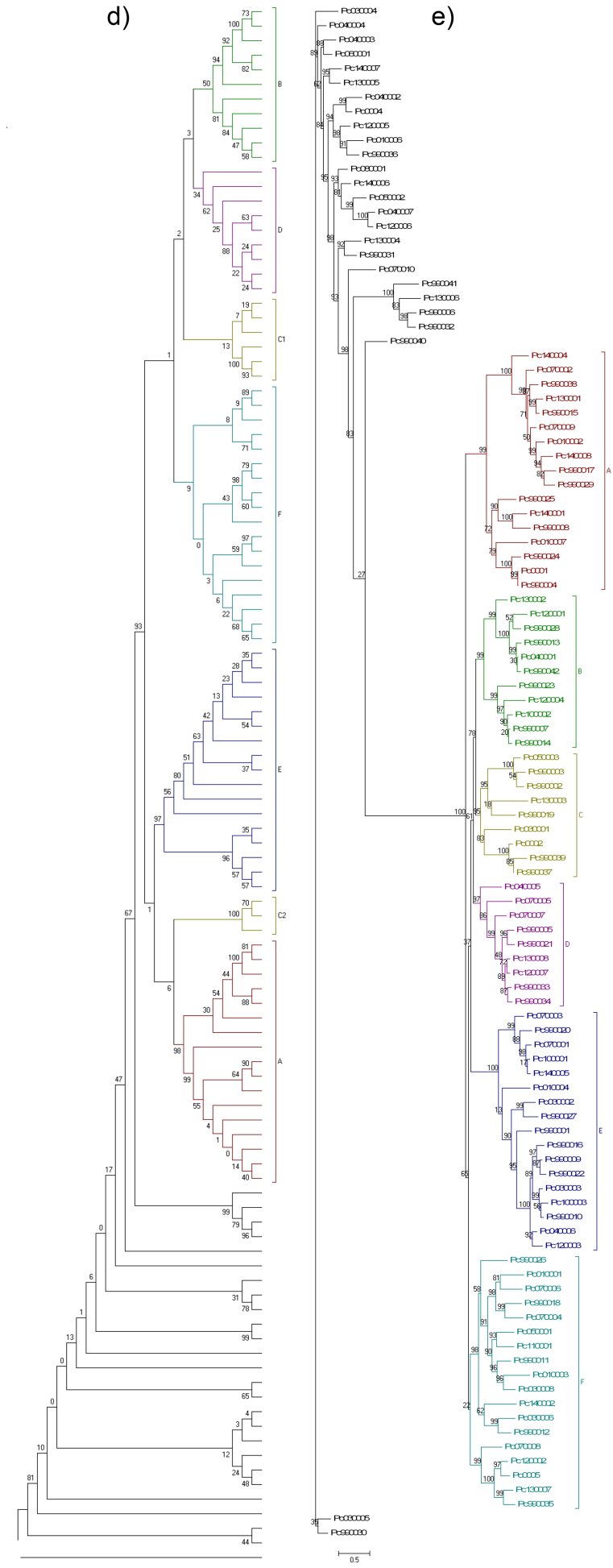


Figure 15: Analysis of the CIR repertoire by different phylogenetic methods.

Phylogenetic trees created using the following methods are shown:

a) un-weighted pair group method with arithmetic mean (UPGMA)

b) neighbour joining

c) minimum evolution

d) maximum parsimony, all implemented by MEGA (Kumar et al., 2004)

e), maximum likelihood, implemented by PhyML (Guindon and Gascuel, 2003, Guindon et al., 2005).

Clades are highlighted with the colour accorded to each subfamily: A) red, B) green, C) yellow, D) pink, E) dark blue, F) light blue and U) black.

The gene names indicated in **e)** were from the first stage of annotation.

An electronic version of this figure is attached in Appendix 2.4.

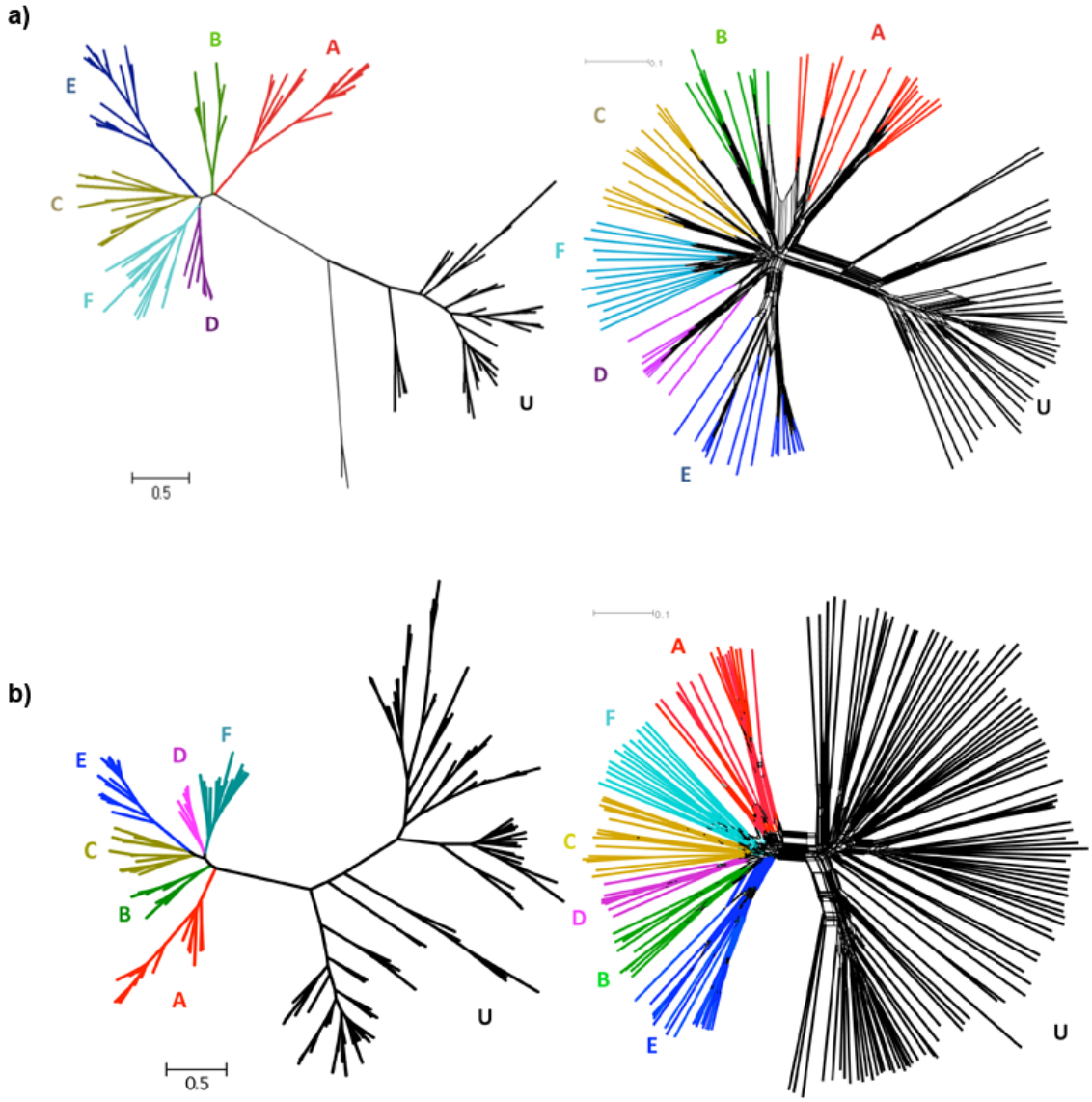


Figure 16: Relationships between CIR amino acid sequences.

Similarities between CIR sequences were visualized using both a Maximum Likelihood tree [left, (Guindon et al., 2005)] and a NeighborNet network [right, (Bryant and Moulton, 2004)].

These phylogenies and networks are shown without branch support values, for ease of view, therefore branches with high confidence [which comprise the majority of branches: defined as >60% calculated by aLRT (Anisimova and Gascuel, 2006) within the phylogenetic trees, and >70% of 1000 bootstrap replicates for networks] are shown in bold, and the figures including branch support values are attached in appendices.

The six sub-families identified using the alignment of 117 CIRs, shown in **a)**, were also seen when more CIR sequences were analyzed, 183 CIRs, shown in **b)**.

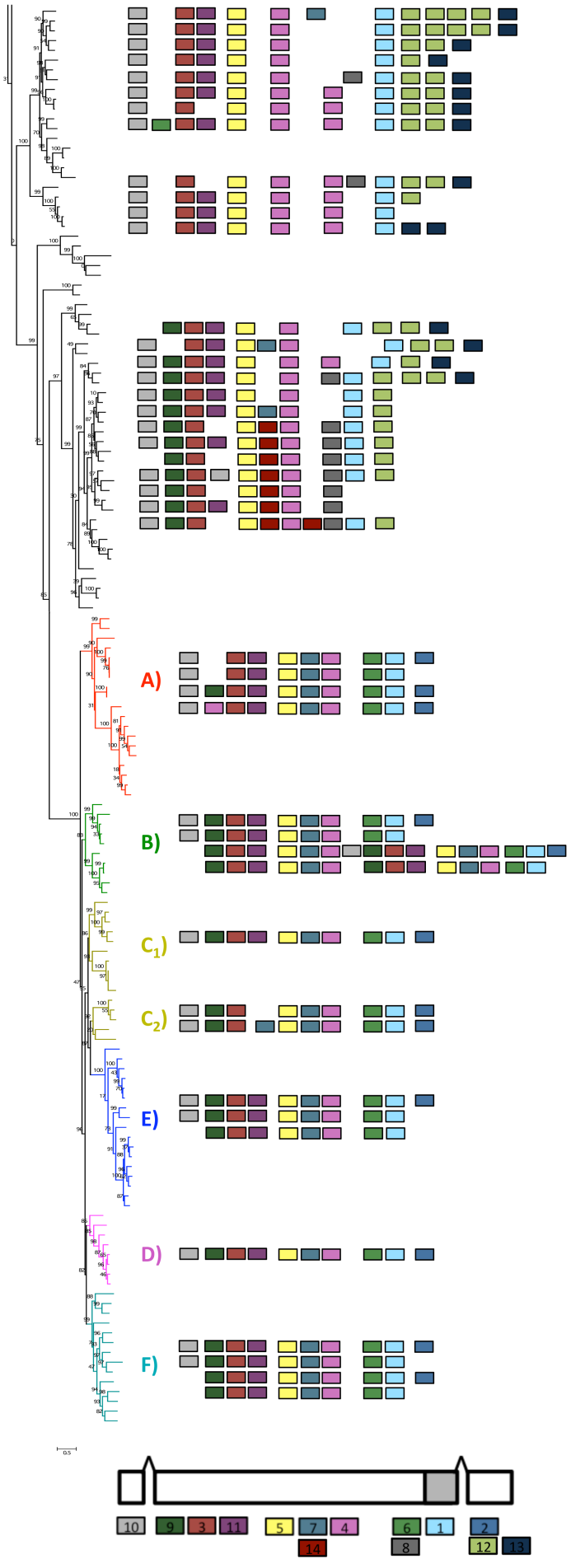

A)
B)
C1)
C2)
E)
D)
F)
10
9
3
11
5
7
4
14
6
1
8
2
12
13

Figure 17: Arrangement of conserved amino acid motifs within CIR sub-families

Amino acid motifs identified by Multiple EM for Motif Elicitation (MEME) analysis [(Bailey and Elkan, 1994), Table 7], were aligned to their position within the encoded protein. A typical CIR is represented at the bottom of the figure, along with the motif key.

To determine whether motifs were specific to CIR subfamilies, they were aligned to the maximum likelihood tree produced using 183 CIRs, Figure 16, which is represented here in a linear fashion. 40 sequences from sub-family U were not analyzed by MEME, this region of the tree is excluded.

The subfamilies A-F are shown in colour at the bottom of the tree: A) red, B) green, C) yellow, D) pink, E) dark blue, F) light blue, whilst subfamily U is black, at the top of the tree.

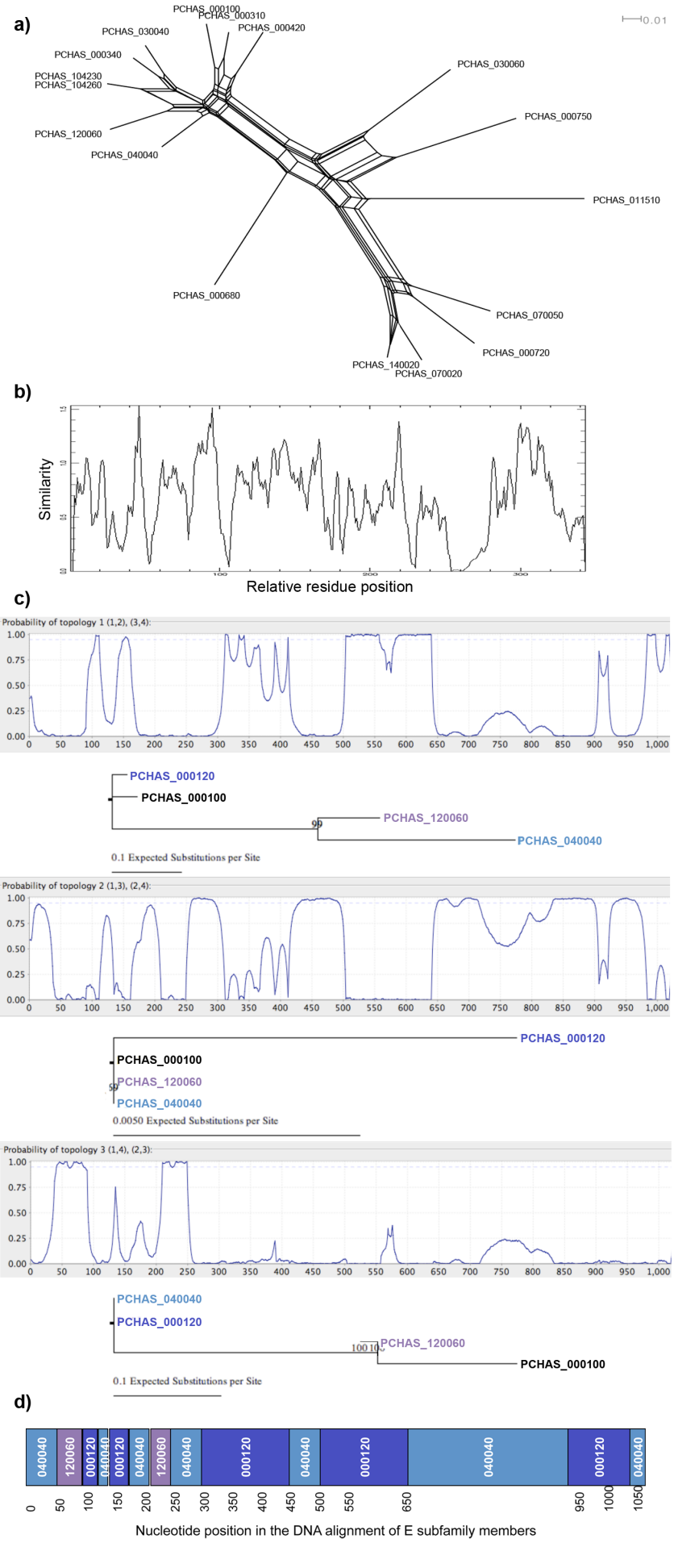

a)
0.01
PCHAS_000100
PCHAS_000310
PCHAS_030040
PCHAS_000420
PCHAS_000340
PCHAS_030060
PCHAS_104230
PCHAS_104260
PCHAS_000750
PCHAS_120060
PCHAS_040040
PCHAS_011510
PCHAS_000680
PCHAS_070050
PCHAS_000720
PCHAS_140020
PCHAS_070020
b)
Similarity
Relative residue position
c)
Probability of topology 1 (1,2), (3,4):
PCHAS_000120
PCHAS_000100
PCHAS_120060
PCHAS_040040
0.1 Expected Substitutions per Site
Probability of topology 2 (1,3), (2,4):
PCHAS_000120
PCHAS_000100
PCHAS_120060
PCHAS_040040
0.0050 Expected Substitutions per Site
Probability of topology 3 (1,4), (2,3):
PCHAS_040040
PCHAS_000120
PCHAS_120060
PCHAS_000100
0.1 Expected Substitutions per Site
d)
Nucleotide position in the DNA alignment of E subfamily members

Figure 18: Evidence for recombination between members of CIR sub-family E.

A NeighborNet network (Bryant and Moulton, 2004), generated using aligned amino acid sequences from CIR sub-family E, **a)**. Sequence identities of the amino acid alignment for CIR sub-family E members were generated in Plotcon (http://emboss.bioinformatics.nl/cgi-bin/emboss/plotcon), **b)**

An example TOPALi profile (Milne et al., 2009, Milne et al., 2004), **c)**, showing the three possible phylogenetic arrangements of sequences PCHAS_000100, PCHAS_000310, PCHAS_120060 and PCHAS_040040 (previously named PCAS_000970, PCAS_000120, PCAS_120060 and PCAS_040040, respectively). The phylogenetic profile indicates how the relatedness of the four analyzed sequences changes along their length. An example phylogenetic reconstruction is found to the right of each profile window, displaying the topology of that relationship.

A reconstruction of PCHAS_000100, **d)**, according to the phylogenetic profile shown above. Each block represents the *cir* sequence which displays closest homology with that region of PCHAS_000100, such that effectively this gene is a mosaic of the other three.

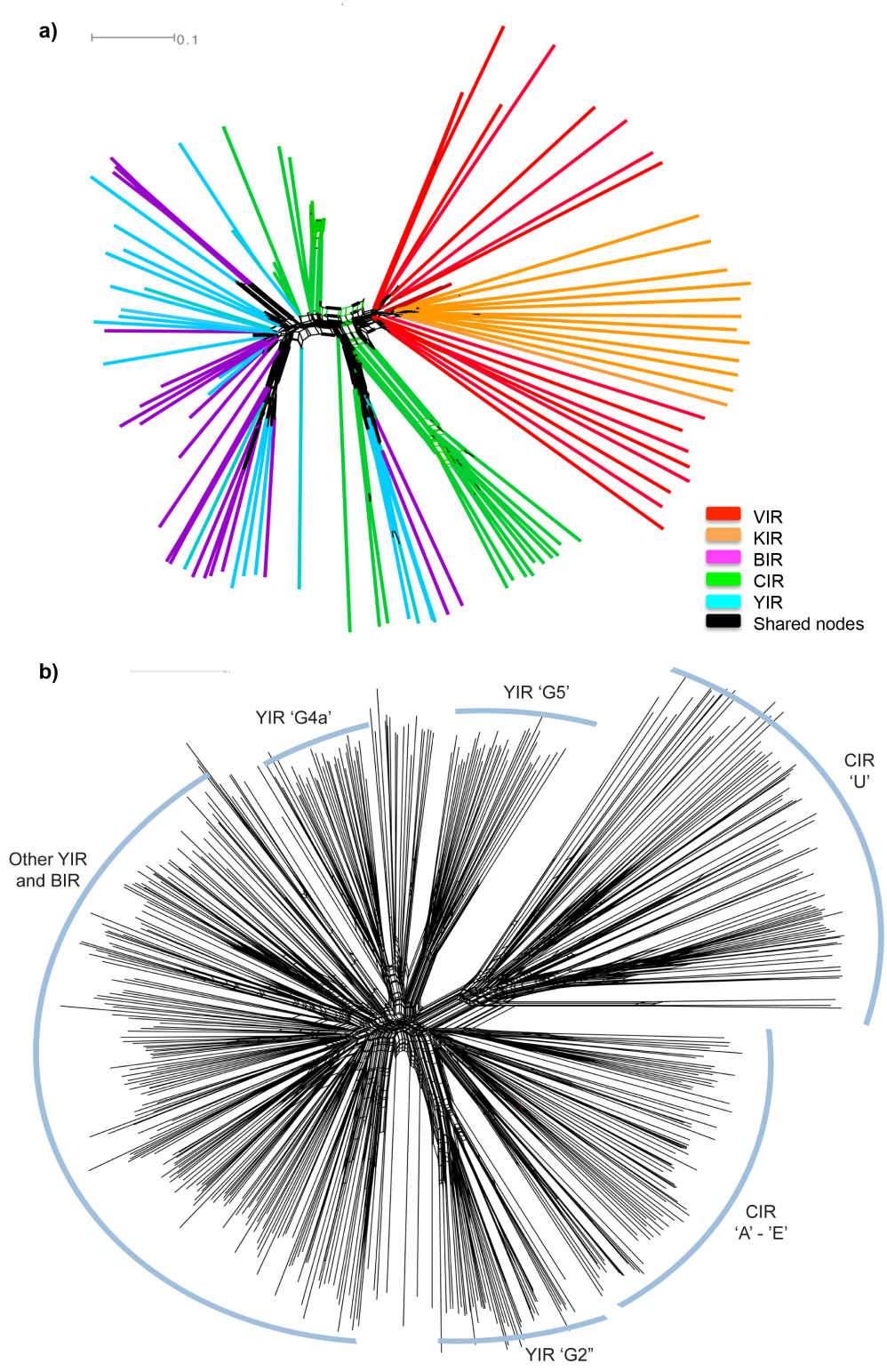


Figure 19: Relations of the CIR repertoire with the PIR super-family.

Adapted from (Cunningham et al., 2010). Networks were created using the NeighborNet algorithm (Bryant and Moulton, 2004), in Splitstree v4.0 (Huson, 1998, Huson and Bryant, 2006). PIR sequences were either downloaded from PlasmoDB (KIR, YIR, & BIR) or kindly provided by Prof. Jane Carlton, New York University (VIR).

136 sequences from the entire PIR superfamily, **a)**, and 500 rodent PIR sequences, **b)**, were compared. The previously identified YIR sub-groups: G2, G4a and G5 (Fonager et al., 2007) and CIR sub-families identified in chapter 2.3.2 are indicated.

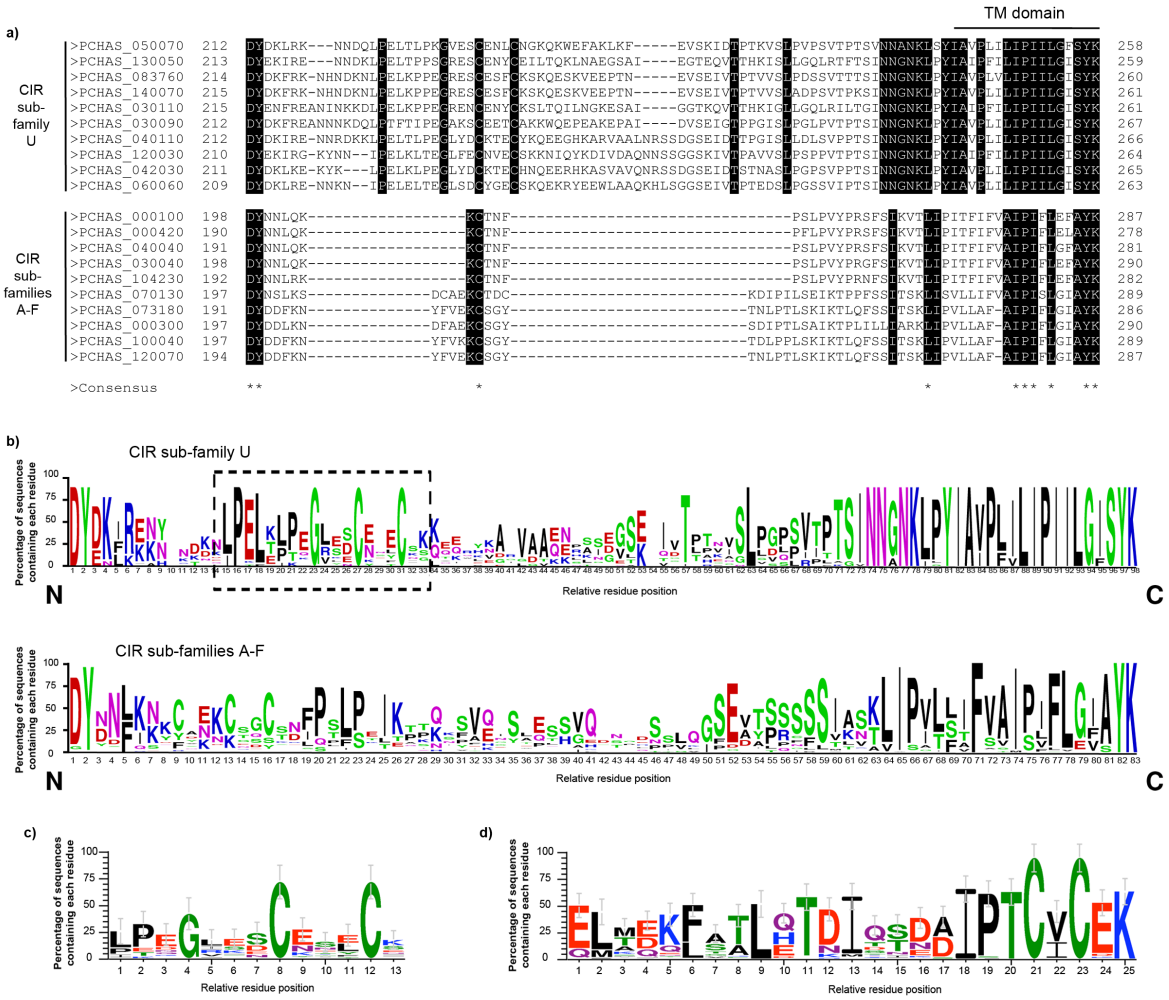


Figure 20: Identification of similarities between the CIR and RIFIN repertoires.

a) Ten representative CIR sequences from subfamilies A-F or U are shown, from an excerpt of the alignment of 183 CIRs. Between the conserved DY and YK residues at the N terminus and C terminus of this region, there are striking differences between these sets of CIRs.

b) A weblogo (Crooks et al., 2004) was created using the sequences in **a).** The most conserved amino acids found in this region of CIR subfamily U, and subfamilies A-F are shown. A striking motif was observed within subfamily U sequences, highlighted by a dotted line.

The subfamily U motif was detected from an alignment of 96 CIR sub-family U sequences, **c)**, and compared to the insert present in the RIFIN subfamily A (Petter et al., 2007), **d)**.

Polar residues are shown in green, basic residues in blue, acidic residues in red and hydrophobic residues in black.

2.4 Discussion

In total, 198 *cir* genes were identified in the *P. chabaudi* AS genome, with the majority being located in sub-telomeric sites and 6% in central locations which could be close to the centromere of those chromosomes. The majority of *cir* genes (96.5 %) retained the typical *pir* gene structure (Del Portillo et al., 2001, Janssen et al., 2002, Janssen et al., 2004, Fonager et al., 2007), consisting of a short first exon, long second exon and highly conserved final exon.

Three long *cir* genes were identified that were approximately double the length of the other *cir*s. These could potentially have been artefacts since DNA sequencing in the sub-telomeres, which are exceptionally AT-rich in *Plasmodium* genomes, is prone to errors (Scherf et al., 2004). For this reason, available transcriptome data was used to confirm the annotation of these *cir* genes. The motif arrangement within these amino acid sequences appears to show a repeated segment, which suggests that these genes were produced by tandem gene duplication events.

Gene duplication and recombination events between members of the *cir* family are likely to have shaped the *cir* repertoire, as observed in other *Plasmodium* multi-gene families including *var* and *sicavar* [for example (Corredor et al., 2004, Frank et al., 2008, Pain et al., 2008)]. Duplication events arise principally from mis-alignment of repetitive DNA, following which the genes evolve in concert by mutation and reciprocal or non-reciprocal recombination, also known as gene conversion. Recombination can either homogenize multi-gene families by promoting shared sequence, or increase polymorphism by the creation of mosaics (Petes et al., 1991, Weill and Reynaud, 1996, Hogstrand and Bohme, 1999, Martinsohn et al., 1999, Kudla et al., 2004, Nikolaidis and Nei, 2004, Maizels, 2005, Jackson, 2007); both mechanisms may act upon the *cir* repertoire.

Fluorescent *in situ* hybridization analyses have demonstrated that chromosome telomeres cluster in the nuclear periphery of *P. falciparum*, enhancing the opportunities for sub-telomeric VSA genes to recombine with one another (Freitas-Junior et al., 2000, Taylor et al., 2000b). Similar localization of telomeres has also been observed in other *Plasmodium* species [reviewed by (Scherf et al., 2008)]. It has been postulated that the frequent recombination allowed by such spatial organization could allow VSA family members to evolve throughout infection, thus maintaining chronicity (Scherf et al., 2008).

Comparison between the genomes of different *P. falciparum* isolates has identified a hierarchy of *var* genes, whereby certain variants appear restricted in their recombination potential (Pain et al., 2008). This suggests that certain *Pf*EMP1 variants are more indispensible than others for maintenance of infection. Attempts were made to identify any *cir* genes similarly exempt from recombination, using comparable methodology to that applied to the *sicavar* genes for identification of recombination events (Pain et al., 2008). No part of the *cir* repertoire appeared exempt from recombination. This could indicate that no such restriction of recombination exists for *cir* genes, as observed for *var*s. More probable though, is that *cir* genes displaying restricted recombination potential were not identified, as the repertoire from only one *P. chabaudi* clone was available for analysis. Whilst bio-informatic measures to identify recombination are crude, particularly within a single genome, they do indicate that recombination has played a strong role in shaping the *cir* repertoire. These observations are consistent with what has been shown in other multi-gene families such the VIR and RIFIN repertoires (Carlton et al., 2008, Joannin et al., 2008).

A crucial question is how frequently recombination occurs between *cir* genes, particularly during mitosis, as the majority of studies using *P. chabaudi* investigate the asexual stages. If recombination occurs frequently in asexual stages, then many analyses of the *cir* repertoire could be meaningless. Attempts to confirm the break-points identified by TOPALi using other methods, such as the recombination detection program [RDP, (Martin and Rybicki, 2000)] were inconsistent (data not shown). This suggests that many of the potential recombination events detected by TOPALi were not recent, and may have been of ancient origin, thus mitotic recombination between *cir* genes may be relatively infrequent. Further analysis of *cir* sequences from other *P. chabaudi* clones and sub-species, including clones that have experienced mosquito-transmission, would be required for a confident estimation of recombination frequency within mitotic and meiotic stages.

The detection of recombination between *cir* family members means that phylogenetic analysis is not an accurate way to analyze PIR sequences as they have not followed a path of linear evolution. Instead, both phylogenetic and non-assumptive methods were employed for this study, to provide a simple overview of the relationships between CIR sequences. Groups of highly similar sequences were identified using information derived from both methodologies. As described above, different tree topologies could arise from different regions of the protein. For example, alignments using only the C-

terminus of CIR sequences may yield very different compatibility results than those using the N terminus. To eliminate this possibility, the whole protein sequences were used for the comparison of relationships between CIRs, with only uninformative regions of the alignments being removed.

Six sub-groups named A-F were initially identified, comprising 75.7% of the 107 CIRs analyzed, whilst the remainder of CIR sequences remained un-clustered. Sub-families have also been identified in the VIR and YIR repertoires. Six VIR sub-families were initially described, and were detected in patient isolates (Del Portillo et al., 2001, Merino et al., 2006). Upon sequencing the whole *P. vivax* genome a further six were identified, meaning that 76% of the VIR repertoire was able to cluster into subfamilies (Carlton et al., 2008, Fonager et al., 2007). Similarly, 56% of the YIR repertoire was able to cluster into five sub-families, 44% of YIRs remaining un-clustered (Fonager et al., 2007). Some of the *yir* sub-families were also detected in *P. berghei*, indicating that, for these closely related parasites at least, some sub-families were shared (Fonager et al., 2007).

In order to detect possible conservation of sub-families, relationships between members of the PIR super-family were compared as described above for the CIRs. PIR proteins were clearly most similar to those from parasites that had undergone the least phylogenetic separation. This supports the original analyses of the PIRs carried out by Janssen and colleagues (Del Portillo et al., 2001, Janssen et al., 2002, Janssen et al., 2004, Fonager et al., 2007). CIR sequences preferentially clustered with themselves, forming two large clades, as seen previously. In particular, sub-family U appeared to be distinct from the other rodent PIRs. These data do not indicate any sharing of sub-families between the *cir* and other *pir* repertoires, and suggest that no sequences are exempt from recombination.

Upon identification of the more divergent *cir* genes by HMM analysis, the proportion of the CIR repertoire represented by the sub-families A-F was reduced to 44.3% of the 183 CIRs analyzed. The remainder of CIR sequences remained un-clustered, comprising 55.7% of the repertoire. Within these sequences, more conserved clades were observed, although due to time constraints these were not investigated separately. Thus the un-clustered group was named sub-family U, and the previously identified sub-groups: A, B, C, D, E and F were defined as the other major sub-family. Given that the two major sub-families each represented almost half of the CIR repertoire, and each contained smaller clades of closely related CIRs, the composition of this family displayed

parallels with that of the RIFINs (Sevier and Kaiser, 2002). Consequently, a more detailed comparison of the CIR and RIFIN families was undertaken.

The RIFIN sub-family A is defined by the presence of a 25 amino acid sequence, which is absent from RIFIN-B types (Petter et al., 2007, Joannin et al., 2008). An insertion was also observed within CIR sub-family U, a section of which contained two highly conserved cysteine residues. The close position of the cysteine residues was similar between the RIFIN and CIR motifs. Cysteine residues often participate in the formation of disulphide bridges [reviewed by (Sevier and Kaiser, 2002)], which suggests that the differential cysteine compositon of CIR and RIFIN sub-families could affect their tertiary structure.

If the members of the CIR and RIFIN sub-families have different protein structures, they could perform distinct functions. This has been predicted to be the case for the A- and B-type RIFINs, using criteria for functional divergence developed using enzymes of known altered substrate specificity (Abhiman and Sonnhammer, 2005, Joannin et al., 2008). This analysis was applied to the CIR repertoire, finding high proportions of rate and conservation shifting sites between the two major sub-families. As for the RIFINs, these data suggest that functional divergence between the CIR sub-families may have occurred.

Notably, no function shift has been predicted for PIR super-family members using the domains curated by Pfam, (http://funshift.sbc.su.se/cgi-bin/stksub.cgi?domain=PF06022&ac=Cir_Bir_Yir&desc=Plasmodium%20variant%20antigen%20protein%20Cir/Yir/Bir). Pfam uses conservative annotation in order to define orthology groups (Sonnhammer et al., 1998, Finn et al., 2010), which means that 86 homology clusters had been defined at the time of writing. Therefore, the definition of two major CIR sub-families by whole-family analysis as presented here may be more informative than identification of many, smaller, highly related clusters of sequences.

In-depth analysis of multi-gene families is time-consuming, particularly when phylogenetic analyses must be carried out in order to identify which sequences are members of each major sub-family. In order to facilitate the identification of A- and B-type RIFINs from field data, a tool has been developed by Joannnin and colleagues: RIFIN subfamily predictor [RSPRED, (Del Portillo et al., 2001, Joannin et al., 2010); available at http:// www.ifm.liu.se/bioinfo/]. This tool was unable to classify CIR sequences within the RIFIN-A or B sub-families. However, if functional similarities

were observed between these protein families, the criteria for classification of CIR sub-types could be included in a similar tool to facilitate their categorization.

Similarities between the CIRs and RIFINs have been previously identified by Janssen and colleagues, who predicted their secondary structures would comprise alpha helices of 12-20 amino acids, separated by coiled-coil regions (Janssen et al., 2004). In addition, *rif* genes were identified by an HMM designed using the rodent *pir* sequences, although at low significance levels (Janssen et al., 2004). Finally, similar amino acid motifs were identified within both repertoires, however these were not consistent across each repertoire.

Motifs have been previously identified in PIR sequences (Janssen et al., 2004, Merino et al., 2006, Carlton et al., 2008). The CIR sequences were searched for evidence of these motifs, but the only motif with detectable similarity was M3 (Janssen et al., 2004), which corresponds to the conserved CIR motif 3. The motifs identified within VIR sequences were not investigated for presence within the CIR repertoire, owing to the lack of detection of other PIR motifs, and the clear divergence of the VIR and CIR repertoires. Pain and colleagues proposed the existence of a PIR protein domain, which may be repeated up to three times beyond the TM domain (Pain et al., 2008). This is likely to also be present within the CIR proteins.

Within the CIR repertoire, 14 conserved amino acid sequence motifs were identified by MEME analysis (Bailey and Elkan, 1994). Sequences similar to the CIR motifs were also detected in a selection of rodent PIR, KIR and VIR sequences. Notably, one of the CIR motifs detected as specific to CIR sub-family U (motif 12) was detected in some rodent PIR sequences. This motif corresponds to the end of the predicted PIR TM domain at the conserved C-terminus.

The majority of PIR proteins are predicted to contain at least one TM domain, a feature which has led to predictions that the C-terminus of most PIR proteins is intra-cellular, whilst the N-terminus is exposed at the iRBC surface (Del Portillo et al., 2001, Del Portillo et al., 2004). Certainly, PIR proteins can be fairly divergent at the N-terminal end, with amino acid identities of only 20-30% between VIR family members (Carlton et al., 2008), supporting the idea that these regions of the proteins are exposed to selective forces. Within a single *cir* gene, evidence for recombination was more pronounced at the 5' end, supporting the idea that recombination here particularly facilitates the generation of diversity. Recombination within the conserved 3' region of

cir genes may be inhibited due to functional constraints, such as the requirement for maintenance of the TM domain. Nonetheless, some CIRs belonging to sub-family U possessed a unique amino acid motif at the beginning of the predicted TM domain, which may give their TM domain different properties to the rest of the CIRs. Possession of a different TM domain may influence the function of these CIRs, perhaps altering which membrane the proteins are able to span.

The lack of typical export sequences questions whether PIR proteins are truly exposed at the iRBC surface. Few PIR proteins contain a predicted signal sequence and a *Plasmodium* export element (PEXEL) / vacuolar transport signal (VTS) motif, which have been shown to be essential for export of certain proteins in *P. falciparum* (Marti et al., 2004, Hiller et al., 2004). The CIR repertoire fits this pattern with only 18 PEXEL motifs detected. However, approximately half of the VIRs contain a PEXEL-like sequence (Pain et al., 2008), and a different possible export motif has been described in the KIR repertoire (Pain et al., 2008). Within the CIR repertoire, 39 PEXEL-like motifs and 2 ZLPS motifs could be detected, found in both major sub-families. It is possible that PIR proteins are exported in a PEXEL-independent manner, which may even require species-specific signals.

Intergenic analyses have been performed already for the *cir* repertoire (Janssen et al., 2002, Janssen et al., 2004). Because VSA genes are known to be flanked by relatively conserved DNA sequences (Voss et al., 2003, Flueck et al., 2010, De Silva et al., 2008), these were not repeated for the present study. Janssen and colleagues suggested that the intron of some *rif* genes shared ancestry with the rodent *pir* genes' second intron (Janssen et al., 2004). However, no homology was observed with the other *pir* intron, or any introns of the *kir* and *vir* genes.

Transcripts derived from the intron act to silence *var* genes, in concert with different nuclear sub-localisation (Voss et al., 2003, Ralph et al., 2005, Duraisingh et al., 2005, Dzikowski et al., 2007), the silent information regulator 2 protein (*Pf*SIR2) and different degrees of chromatin modification [reviewed by (Dzikowski and Deitsch, 2009, Scherf et al., 2008)]. It is possible that this could also be the case for *pir* genes, as a conserved motif has been identified in *vir* introns, near the donor splice site (Janssen et al., 2004, Merino et al., 2006, Carlton et al., 2008). *P. vivax* telomeres also localize to distinct regions of the nuclear periphery (Scherf et al., 2004), and *yir* genes were found to have different intronic types (Fonager et al., 2007). If the second intron is involved in

elements of *pir* regulation, then *rif* gene expression could also be controlled by the same mechanism due to the conservation found between their introns.

In previous studies, motifs have been identified outside of the *pir* coding sequence, implying there are different aspects of regulatory control for *pir* genes. A motif was identified upstream of *yir* genes (Fonager et al., 2007), which is present in all rodent *pir* promoters. This motif could not play a role in differential expression of individual *pir* genes, but could act perhaps as a transcription enhancer. A brief search of the regions 500 nucleotides upstream and downstream of *cir* genes identified some motifs with similarity to those identified for *yir*s, such as the M2 motif found in 58% of *yir*s and 43% of *cir*s (Fonager et al., 2007).

Recently, the discovery of the first class of *Plasmodium* transcription factors (having an Apetala 'AP2' domain similar to the plant AP2/ERF DNA-binding proteins (Balaji et al., 2005), has enabled mapping of some of their substrates (Voss et al., 2003, Flueck et al., 2010, De Silva et al., 2008). Some of these sites could also be detected in *cir* intergenic regions (Table 9), indicating that *cir* transcription could in part be regulated by binding of AP2 transcription factors, as proposed for members of the upsB and upsC *var* gene subfamilies by the AP2 transcription factor PF14_0633 (De Silva et al., 2008), and *Pf*SIP2, which has been shown to be involved in *var* gene silencing (Voss et al., 2003, Flueck et al., 2010, De Silva et al., 2008)

Table 9: Detection of *Plasmodium* ApiAP2 transcription factor DNA binding sites upstream of *cir* genes.

Feature	Sequence	Proportion of *cir*	Reference
PF14_0633 (*P. falciparum*) Active in intra-erythrocytic stages	CATGC	15.7%	(Yuda et al., 2009)
AP2-O (PB000572.01.0 *P. berghei* /PF11_0442 *P. falciparum*) Active in ookinete stages	TAGCTA	38.6%	(De Silva et al., 2008)
*Pf*SIP2 (PFF0200c *P. falciparum*) Active in intra-erythrocytic stages	GTGCA [GCA]TGCA & [ACG]GTGC[GA] *	22.8% 100% & 29.1%	(De Silva et al., 2008)

* *Pf*SIP2 has a bipartite DNA binding site (Flueck et al., 2010), both sequences were searched independently in the *cir* upstream regions.

As yet the significance of the identified CIR sub-families and amino acid or potential regulatory motifs are unknown. Microarray analyses have previously shown no correlation between phylogenetic sub-group and *pir* transcription (Bozdech et al., 2008, Cunningham et al., 2009). However, the similarities observed here between the two

major CIR sub-families and the RIFINs, suggest that the CIRs may display differential expression and localization as has been shown for the A- and B- type RIFINs (Petter et al., 2007).

The data presented here provide an ideal starting point for experimental investigation of the *cir* family, including association of potential regulatory motifs with their expression, and validation of the identified sub-families. The identification of similar CIR sequences, and conserved amino acid motifs inform the design of reagents to carry out such investigations, which are presented in this thesis as follows:

- Based upon the *cir* gene annotation presented here, primers were designed for the measurement of *cir* transcription by quantitative reverse-transcription PCR, described in chapter 3.
- The amino acid sequences of particular *cir* genes were chosen for expression as recombinant proteins, described in chapter 4.
- Synthetic peptides corresponding to conserved and sub-family specific amino acid motifs were synthesized, and anti-sera raised to these, which were used for detection of CIR proteins in iRBC, described in chapter 5.
- In addition, the antibody responses to these peptides were measured in *P. chabaudi* infected mice, described in chapter 6.

The repertoires of different *P. chabaudi* AS clones, and different *P. chabaudi* strains should be compared, to enable detection of changes such as recombination events, and their frequency. In addition, efforts should be made to determine the extent of recombination occurring between *cir* genes during meiosis, which would require mosquito transmission of the parasite.

Since frequent recombination events could potentially confound further investigations of these genes, knowledge of the recombination frequency, and if indeed any *cir* are exempt, would allow better experimental design. The extensive diversity in VSA genes observed in the field may also be better understood by detailed analysis of the variant gene families present in laboratory models such as *P. chabaudi* AS.

Little is known about the extent of *vir* sequence diversity, short of the amplification by PCR of particular sub-family members (Del Portillo et al., 2001, Joannin et al., 2010). In *P. falciparum* the VSA repertoires are extremely diverse, with different copy numbers of *rif* genes existing in different parasite isolates (Kraemer et al., 2007) and only three *var* genes which are consistently shared between isolates [for example: (Bull

et al., 2008)]. An approach taken by Bull and colleagues was to create networks using *var* gene sequence 'tags' from different patient isolates (Bull et al., 2008). By comparing whole *var* gene sets between different parasite populations, they were able to find blocks of highly polymorphic sequence, shared between different *var* genes. By connecting *var* genes which contained these blocks, and connecting these to *var* with similar sequences, a network was created that spanned the majority of the *var* sequence tags used (Bull et al., 2008). The authors identified putative groups of *var* genes which have recombined, and suggest that the *var* genes have mosaic structures due to the problem of maximising antigenic diversity, minimising epitope sharing between variants whilst maintaining the protein binding function. Such an analysis could be feasible in the future for *vir* genes, as sequencing data from field studies accumulate. The mechanisms of *pir* repertoire structuring could be supported by dissection of the *cir* genes in the rodent model *P. chabaudi*, under known selective conditions.

Chapter 3: Analysis of *cir* transcription

3.1 Introduction

Characterization of the *cir* gene family, described in chapter 2, raised the question of which *cir*s were transcribed during the blood stages of *P. chabaudi* infection. Several studies have investigated *pir* transcription during blood stages of the *Plasmodium* lifecycle. *In vitro* analysis of *vir* expression during the 48h erythrocytic growth cycle has indicated that at least 59% of *vir*s were transcribed (Bozdech et al., 2008, Cunningham et al., 2009). Similarly, 42% of the larger *yir* family was transcribed *in vivo*, during infection of immuno-deficient mice (Cunningham et al., 2009). Different *yir*s were more highly transcribed in ring / trophozoite stages and in schizonts (Cunningham et al., 2009), supporting the observation of two waves of *vir* and *Pvtrag* transcription detected during *P. vivax* intra-erythrocytic development (Bozdech et al., 2008, Cunningham et al., 2009). These studies were carried out using microarrays, which allow an insight into the transcription of the whole multi-gene family at a given time.

More recently, with the advent of highly efficient DNA sequencing technologies, it has become feasible to apply genome level analyses to the transcriptome, by methodologies known as RNA sequencing. This technique has several benefits compared to the use of microarrays for measuring the expression of multi-gene family members, particularly since the design of gene specific probes is not required and data may be re-analyzed upon improvements to the genome annotation status. RNA sequencing has been used in a pioneering study of the *P. falciparum* transcriptome, comparing different time-points within the intra-erythrocytic development cycle (Otto et al., 2010). Not only did this methodology confirm results from microarrays performed in tandem, but more low abundance transcripts and other gene features could also be detected including splice variants. The application of this technology to other *Plasmodium* species would be extremely useful, particularly for the measurement of multi-gene family transcription.

Reverse transcription followed by quantitative PCR has also been used to address *yir* transcription: to validate the expression of individual *yir* genes detected by microarray (Cunningham et al., 2009, Shi et al., 2005), and to detect transcription of *yir* sub-family members (Fonager et al., 2007). This technique enables the detection of *pir* transcripts from small quantities of RNA, even from individual iRBCs (Fernandez-Becerra et al.,

2005). Individual *P. yoelii* iRBCs have thus been shown to transcribe 1-3 *yir* transcripts (Cunningham et al., 2009), whilst members of at least two *vir* sub-families have been detected from individual *P. vivax* iRBCs (Fernandez-Becerra et al., 2005).

The objectives of the work described in this chapter were to:

i) Design primers within *cir* genes and to validate their specificity and utility for RT-qPCR analyses.

ii) Measure *cir* transcription during *P. chabaudi* infection by RT-qPCR.

iii) Investigate whether whole transcriptome RNA sequencing analysis could be applied for detection of the whole *cir* repertoire during *P. chabaudi* infection.

3.2 Methods

3.2.1 Mice and parasites

All experiments used female BALB/c mice, using either wild-type mice or mice with a targeted disruption in the *RAG2* gene [$RAG2^{-/-}$, (Shinkaia et al., 1992)]. Mice were of 5 - 8 weeks of age at the time of experiment initiation and were kept in living conditions according to the Home Office animal act (1980) regulations.

P. chabaudi AS parasites were cloned and maintained at the National Institute for Medical Research (Slade and Langhorne, 1989), with only four passages in wild-type mice from the original isolate provided by Professor D. Walliker (University of Edinburgh). As further cloning would give rise to stochastically different populations of parasites following their expansion, all infections were initiated using *P. chabaudi* AS expanded in $RAG2^{-/-}$ mice from a reference stabilate produced by limited serial passage in the same mouse strain. This procedure was carried out such that all infections could be initiated with a highly similar parasite population, since the passage of parasites within immuno-competent mice was expected to place them under lymphocytic selection (Cunningham et al., 2005).

All infections were initiated by intraperitoneal (i/p) injection with *P. c. chabaudi* AS iRBC. Blood samples were taken from mice either under terminal anaesthesia, for collection of the total blood volume by cardiac puncture, or by removal of the tail tip, for samples less than 100 µl. Blood samples were taken into Krebs saline (114 mM NaCl, 4.57 mM KCl, 1.15 mM $MgSO_4$ (Krebs and Eggleston, 1940), containing 0.2% glucose and 25 U/ml heparin (Leo Pharmaceuticals). Blood was either used immediately for flow cytometry, reticulocyte depletion experiments, to make thin blood films, or stored at -80 °C in TRIZOL reagent (Invitrogen) for subsequent RNA extraction.

3.2.2 Preparation and counting of thin blood films

Parasitaemia was monitored by daily microscopic analysis of methanol-fixed Giemsa-stained thin blood smears. Standard thin blood films were prepared using 2 µl blood, which was drawn across the surface of superfrost-coated slides (Menzel Glazer), allowed to air dry, and fixed using methanol. After fixed slides were dry, they were

stained using Giemsa (VWR international), diluted to 2x concentration in buffer (comprising 20mM Na_2HPO_4 and 40mM KH_2PO_4, pH 7.4).

For measurement of reticulocytes, Brilliant Cresyl Blue slides were prepared by coating the slide surface with 0.3% Brilliant Cresyl Blue (BDH) dissolved in absolute ethanol. Once dry, blood films were drawn across the surface of these slides, and incubated in a moist chamber for 15 minutes to allow RBCs to absorb the stain. Subsequently, these slides were allowed to dry, fixed and Giemsa stained, as standard thin blood films. This method has been adapted from previously described work (Zuckerman, 1957).

Slides were counted using a Zeiss Axioplan microscope, with a 25x eyepiece and 100x objective lens (Zeiss), under oil immersion (Zeiss). At low parasitaemia, the number of RBCs in each field was enumerated, and the number of iRBCs was recorded in each field until an estimated 10,000 RBCs were counted. When parasitaemia surpassed 5 % of RBCs infected, the number of iRBCs in a representative field of 200 RBCs was recorded. Enumeration of reticulocytes was carried out as a proportion of total RBCs, as described for iRBC counts at low parasitaemia (ie the number of reticulocytes present in approximately 10,000 RBCs).

3.2.3 Primer design

Primers were designed to amplify 150 – 250 nucleotides within selected *cir* genes using Primer3 software (Rozen and Skaletsky, 2000), with default settings except: 18 - 30 nucleotide primer length and 55 - 65 °C primer melting temperature. Two pairs of primers were designed in the same way to amplify *Mus musculus* haemoglobin beta chain, to determine the extent of host RNA contamination in extracted parasite RNA for full transcriptome sequencing. 36 primer pairs were synthesized by Sigma Aldrich, listed in Appendix 3.1.

3.2.4 Endpoint-polymerase chain reaction

Primer efficacy was screened via endpoint PCR, using genomic DNA as the template. One µl each of forward and reverse primers were used at 10 pMol / µl concentration, in the presence of 2.5 µl 10x buffer, 1.5 µl 25 mM magnesium chloride, 1.25 units Taq polymerase (New England Biosciences), 0.1 µl 10 mM dNTPs, 50 ng gDNA template,

and sterile water to 25 µl final volume. Amplification conditions were 94 °C for 2 minutes, followed by 35 cycles of: 94 °C for 1 minute, 65 °C for 1 minute and 72 °C for 1 minute, with a final extension of 72 °C for 7 minutes, before cooling to 4 °C. Amplified products were resolved on a GelRed stained (Biotium), 2% agarose gel for 30 minutes at 90 V. Product sizes were resolved using hyperladder IV (Invitrogen). These data are attached in Appendix 3.2.

3.2.5 Primer specificity determination

Primer specificity to the intended *cir* genes was determined by direct ligation of PCR-amplified products into the TA cloning vector pCR2.1 (Invitrogen) followed by transformation into competent *E. coli*, according to manufacturer's instructions. Briefly, the ligation reaction used a 3:1 ratio of fresh PCR product diluted in sterile water (3 ng) to vector (25 ng), in a final volume of 10 µl in the presence of 1 µl of 10x reaction buffer and 400 units of T4 ligase (New England Biosciences). Ligation reactions were incubated overnight at 14 °C.

Two *Escherichia coli* strains were used for transformation, first DH5α (Sigma Aldrich), and secondly *Epicurian coli* gold super efficient cells (Stratagene). For DH5α cells, 15 µl cell aliquots were thawed on ice to which 2 µl ligation reaction was added, and incubated for a further 10 minutes on ice. Cells were heat-shocked for 30 seconds at 42 °C, allowed to recover on ice, before 100 µl SOC medium (Invitrogen) was added and cells incubated, shaking at 37 °C for one hour. 50 µl cells were then plated onto Luria Bertani (LB)-agar plates containing 50 µg/ml ampicillin and 1 mg/ml x-galactose (both Sigma) and incubated overnight at 37 °C. The super efficient cells followed the same protocol as the DH5α cells, except for a 10 minute incubation on ice with of 1 µl β mercaptoethanol solution (Stratagene) before addition of the ligation reaction, and the requirement of 100 µM Isopropyl β-D-1-thiogalactopyranoside (IPTG) presence in the LB-agar plates to allow blue-white screening.

Between three and five transformants were selected, which were cultured overnight in LB broth containing 50 µg / ml ampicillin (Invitrogen), and plasmid DNA was extracted via miniprep (Qiagen), according to manufacturers instructions. Plasmid inserts were sequenced using the M13 forward primer (Source Biosciences), and sequences were verified by basic local alignment tool searches [BLAST, (Altschul et al., 1990)] against

the *cir* database (curated at http://www.genedb.org/blast/submitblast/GeneDB_Pchabaudi, by the Wellcome Trust Sanger Institute).

3.2.6 RNA extraction

Blood stored at -80 °C in at least 10 volumes of TRIZOL reagent (Invitrogen) was defrosted on ice, and 40% of the TRIZOL volume of chloroform: isoamyl alcohol (24:1; Sigma Aldrich) was added. After incubation on ice for five minutes, samples were centrifuged at 13,000 x *g* for 30 minutes at 4 °C and the aqueous phase removed. RNA was precipitated by addition of an equal volume of isopropanol (Invitrogen) and incubation on ice for one hour. Samples were centrifuged as before, and the RNA pellet washed with 70% ethanol (Invitrogen). Samples were centrifuged as before, supernatants removed and the RNA pellets air-dried, inverted, for five minutes before re-suspension in DEPC-treated water (Invitrogen). Nucleic acid concentration was estimated by spectrophotometry (Nanodrop; ThermoFisher).

500 ng RNA was DNAse digested, by addition of 2 µl Turbo DNAse buffer and 1 µl Turbo DNAse (both Ambion) and sterile water added to a final volume of 17 µl. Samples were heated at 37 °C for 30 min, and the reaction terminated by addition of 2 µl Turbo DNAse inactivation reagent (Ambion). Samples were mixed gently, centrifuged at 10,000 x *g* for 2 minutes, and the RNA transferred to fresh tubes, before storage at -80 °C.

All RNA concentrations were estimated using spectrophotometry (NanoDrop, ThermoScientific) after extraction via the phenol-chloroform method (Sambrook and Russell, 2001). The ratio of the absorbance at 260 and 280nm ($A_{260/280}$) was determined (Appendix 3.3), which allows estimation of nucleic acid purity, whereby un-contaminated RNA should have an $A_{260/280}$ approximately equal to 2. The ratio for pure DNA should be 1.8, therefore RNA samples contaminated with DNA would have $A_{260/280}$ ratios closer to 1.8 than 2. Lower $A_{260/280}$ ratios could also be attributed to protein contamination, since amino acids absorb strongly at 280 nm (Bustin et al., 2009). However there are limitations to the extent which spectroscopy can determine nucleic acid quality, therefore formamide-gel electrophoresis was also performed on all RNA samples to ensure the absence of degraded material (Appendix 3.3).

3.2.7 Reverse transcription

Reverse transcription uses primers within RNA sequences to generate the complementary DNA sequence (cDNA), a more stable template for subsequent investigation. All cDNA samples used for estimation of *cir* and *beta globin* transcription throughout this thesis were reverse transcribed at the same time. This was carried out using 250 ng of DNAse digested RNA, to which 0.52 µg random hexamers (Promega), 1 µl 10 mM dNTPs (Invitrogen) and sterile water were added to a final volume of 12 µl. Samples were heated at 65 °C for 5 minutes before addition of: 4 µl first strand synthesis buffer, 1 µl 0.1 M dithiothreitol (DTT), 0.5 µl RNAsin and either 0.5 µl Superscript II (Invitrogen) for RT positive or 0.5 µl H_2O for RT negative controls. Samples were heated at 42 °C for 50 min, followed by 70 °C for 15 min, and 4 °C overnight, before storage at -20 °C.

3.2.8 RT-qPCR

Quantitative PCR uses fluorescent dyes or sequence tags to detect PCR products as they accumulate throughout the reaction. The principle of this technique is that fluorescence intensity is measured on a logarithmic scale against cycle number as the PCR reaction progresses. A threshold value is set such that background fluorescence is excluded, and the cycle when sample fluorescence reaches the threshold is recorded as the quantification cycle (Cq). The design of RT-qPCR experiments was carried out according to the minimum information for the publication of quantitative PCR (Bustin et al., 2009), guidelines which attempt to standardize the reporting and design of RT-qPCR experiments, including the use of a universal nomenclature.

All quantitative PCR was carried out using 12.5 µl SYBR green ROX mastermix (Fermentas) with primers at 0.2 µM concentration, in a final volume of 25 µl. All reactions were carried out in triplicate, with every primer containing no template control and positive control (gDNA for *cir* primers, cDNA pool for *beta tubulin* primers) samples on each plate.

Primer sensitivity was determined using gDNA template titration, diluted in sterile water (Gibco) to give 0.5 ng, 1 ng or 5 ng DNA per 25 µl total reaction volume. Thereafter RT-qPCR was carried out using all samples at 5 ng cDNA per 25 µl total reaction volume, with all RT- and positive controls at the same concentration.

Quantitative PCR conditions were 50 °C for 2 minutes, 95 °C for 15 minutes, followed by 35 cycles of 95 °C for 15 seconds and 60 °C for 1 minute. Because all samples were quantified in relation to internal plate controls, two qPCR machines were used in parallel: the 7000, using SDS 2.0 software and the 7900, using SDS 2.0 and 2.2.2 software (both ABI). Quantification cycle numbers (C_q) were determined using the automatic threshold level of 0.2 within the SDS software, then all C_q values were exported into excel for further analysis (all data analysis is contained in Appendix 3.4). To validate the entire plate, samples were only analyzed further if the no template control (NTC) and RT- samples gave the maximal Cq values, and if the positive controls (gDNA / cDNA reference sample) were similar to those measured previously.

3.2.8 i) Analysis of **cir** *transcription*

Nucleic acid quantification can be achieved either by comparison with standards containing known amounts of the target gene, known as absolute quantification or by relative quantification, where the differences between samples are compared, after normalization to a control constitutively expressed gene. As no standards were available, the latter approach was used here. *cir* gene expression was normalized to another *P. chabaudi* reference gene, *beta tubulin*, and analyzed using the efficiency corrected ratio method (Pfaffl, 2001). Such normalization permits comparison between samples whose RNA concentration, efficiency of reverse transcription and amplification efficiencies are slightly different, assuming that the reference gene is expressed equally in the different samples (Bookout et al., 2006).

Beta tubulin primers have been described elsewhere: AGCAGGCCAATGTGGTAATC, Forward; ACCTGCACGAACACTATCCA, Reverse (Cunningham et al., 2009).

For each sample, the mean and standard deviation (SD) of the three technical replicates was measured (Raw data and analyses are contained in Appendix 3.4). Outlier points were discarded according to standard methodology (Bookout et al., 2006). Samples were analyzed where the confidence values were 99%, 95%, 90%, 85% and 80% (SDs of 0.3, 0.4, 0.525, 0.650 and 0.775, respectively). Samples were excluded if the confidence value was less than 80%.

PCR efficiency was calculated using the logarithm of the gDNA amounts used in a titration experiment, measured by qPCR. The linear regression was determined from

these data, and slope values inputted into the equation: $E=10^{(-1/slope)}$ (Bookout et al., 2006). However, since the *beta tubulin* primers spanned an exon: exon boundary, the product size from a gDNA template was too large to be efficiently amplified by qPCR. For this reason, comparison of PCR efficiencies between these reference primers and the *cir* primers required cDNA template titration.

The ratios of the Cq difference between the *cir* transcripts present in different samples, normalized to the reference gene *beta tubulin*, were calculated according to the method described by Pfaffl (Pfaffl, 2001). This methodology was used because the *cir* primers did not have equal efficiencies of amplification with the reference primers for *beta tubulin.* The following equation was used to calculate the ratio of *cir* transcripts, where E is the PCR efficiency for each primer pair and ΔCq refers to the Cq difference between the two samples under comparison:

$$\text{Ratio} = (E_{cir})^{\Delta Cq \text{ of } cir \text{ (control-sample)}} / (E_{beta\ tubulin})^{\Delta Cq \text{ of } beta\ tubulin \text{ (control-sample)}}$$

The standard deviation (SD) for the ratio of *cir* transcripts detected, relative to the control sample, was determined using the following equation:

$$SD\ \Delta Cq = \sqrt{[(SD_{beta\ tubulin})^2 + (SD_{cir})^2]}$$

3.2.8 ii) Analysis of beta globin *mRNA contamination*

A measure of *M. musculus* contamination of *P. chabaudi* mRNA was defined using the mean Cq values from the three technical replicates. The reciprocal Cq for *beta globin* was calculated as a percentage of the reciprocal Cq of *beta tubulin* to determine whether *beta globin* mRNA levels were higher in the sample than those of *P. chabaudi* mRNA. No correction for differences in qPCR amplification efficiency was performed here, thus the detected *M. musculus* mRNA contamination may in fact be higher than indicated by the ratios, as the *beta globin* primers had a lower amplification efficiency than the *beta tubulin* primers. Nonetheless, this method serves as a simple indicator of whether *P. chabaudi* mRNA samples contained substantial host contamination. Standard deviation was calculated using the equation above. All related information is attached in Appendix 3.6.

3.2.9 RNA sequencing

Blood of a *P. chabaudi* AS infected BALB/c mouse on day 7 of infection was collected as described above (section 3.2.1), diluted 15-fold in PBS prior to filtration through commercially available Plasmodipur filters (Euro Diagnostica), and the RNA extracted. Two samples were prepared in this way, from separate infections. For cDNA synthesis, a combination of random oligonucleotide and oligo(dT) primers was used as previously published (Bozdech *et al.*, 2003), followed by treatment with Terminator™ 5′-Phosphate-Dependent Exonuclease (Epicentre) for removal of un-5' capped mRNA, ribosomal RNA and transfer RNA species, following manufacturer's recommendations.

P. chabaudi AS RNA samples were prepared according to the methods described by Otto and colleagues (Otto et al., 2010). Firstly, a cDNA library was created by nebulisation (35 psi for 6 minutes) to shear the template, followed by end-repair using T4 polynucleotide kinase to blunt-end the DNA fragments, Klenow polymerase and T4 DNA polymerase. A single 3′ adenosine was added to the cDNA using Klenow exo-polymerase and dATP.

Complementary DNA library fragments were ligated to adaptor oligonucleotides, containing primer sites for sequencing and flowcell surface annealing. Gel-electrophoresis was used to separate library DNA fragments of 200-250 bp in size from unligated adapters, followed by gel extraction at room temperature to avoid under-representation of AT-rich sequences (Quail et al., 2008). The QIAquick kit (Qiagen) was used for this purpose, according to manufacturer's instructions. Hybridization of the adaptors to complementary sequences attached to the flowcell surface used an iso-thermal DNA polymerase to amplify bound DNA fragments into clusters of identical molecules on the flowcell surface. Libraries were amplified by 18 cycles of PCR with Phusion DNA polymerase (Finnzymes Reagents), for sample 1, and Herculase II DNA polymerase (TwistDX), for sample 2. This process is known as solid phase bridge amplification, and can create up to 1000 identical copies of each DNA molecule (Technology Spotlight: Illumina® Sequencing, http://www.illumina.com/support/literature.ilmn).

The Illumina GA II platform, initially developed by Solexa, is based on the sequencing-by synthesis reaction. Sequencing-by-synthesis incorporates a reversible terminator nucleotide, each labeled with a different fluorophore fluorescently labeled, at every step

of the sequencing reaction. At each step only one base may be incorporated and subsequent imaging enables determination of which nucleotide bound. The 3' end of the nucleotides may then be un-blocked for the next sequencing step to occur. Following completion of the synthesis reaction, the sequence of each cluster may be determined by analysis of the images. *P. chabaudi* sequencing libraries were diluted to 2 nM, denatured with sodium hydroxide and diluted to 3.5 pM in hybridisation buffer (Illumina). Following this, libraries were loaded onto a single lane of an Illumina GA flowcell. Cluster formation, primer hybridisation and sequencing reads from 54 cycles were performed using the manufacturer's recommended reagents and protocol (Illumina, https://icom.illumina.com/).

The programs Sequence Search and Alignment by Hashing Algorithm 2 [SSAHA2, (Ning et al., 2001)] and SSAHA_ pileup were used to align the Illumina reads against the *P. chabaudi* AS reference genome curated at http://www.genedb.org/Homepage/Pchabaudi. During SSAHA2 mapping, reads were only included if one end of the pair aligned uniquely to the genome and the distance between the pairs was within the expected insert size. The output of SSAHA_ pileup was used to create the coverage plots over the genome. Uniqueness plots for all possible windows of 50 base pairs (bp) over the genome were generated, and repeat regions longer than the gene read length were ignored for the expression calculation. The expression levels of each gene were calculated based on the geometric mean depth of sequencing coverage of the gene.

These data are attached in Appendix 3.5.

3.2.10 Flow cytometry

P. chabaudi AS iRBCs from RAG2$^{-/-}$ and wild-type BALB/c mice were washed three times in PBS, using centrifugation at 1500 x*g* at 4 ˚C for 5 minutes. Total cell counts were performed using a haemocytometer (NeuBauer), and blood diluted accordingly to give a final concentration of 1x10^7 RBCs per ml in flow cytometry buffer (PBS containing 1% BSA (w/v), 2 mM ethylene-diamine-tetra-acetic acid (EDTA) and 0.01% sodium azide).

The RBCs (1x10^6) were plated in triplicate into V-bottom plates (Nunc), centrifuged at 1200 x *g* for 1 minute at 4 ˚C and the supernatant removed. Cells were then stained with a combination of two dyes, of which thiazole orange (Sigma-Aldrich) binds both DNA

and RNA, whilst DRAQ5 (Biostatus Limited) binds DNA only. Thiazole orange (Sigma-Aldrich) was used at 1 mg/ml and DRAQ5 (Biostatus Limited) was used at 50 μM, diluted in blocking buffer, and incubations performed in a volume of 100 μl for 10 minutes on ice. Following this incubation step, cells were re-suspended twice in 150 μl blocking buffer, followed by centrifugation at 1200 x *g* for 1 minute at 4 ˚C. Cells were re-suspended in 150 μl flow cytometry buffer for data acquisition.

Single colour and unstained controls were also performed for all samples, to allow calibration of the flow cytometer. Immediately prior to data acquisition, samples were passed through a 45 μm filter (Nunc), to ensure no aggregates were present. At least 300,000 events were acquired within a defined region, which excluded very small or dead cells. Data were acquired using the CyAN flow cytometer (Dako Cytomation), using default filter settings, with Summit™ software (Dako Cytomation). Following data collection, information was further analyzed using FlowJo software version 8.8.6 (Tree Star). A region including all live RBC was drawn in FlowJo software (version 8.8.6.2), followed by a region excluding DNA positive cells (DRAQ5+). As mature reticulocytes are the only cell type to be RNA positive but DNA negative, remaining thiazole orange positive cells were enumerated.

3.2.11 Reticulocyte depletion by magnetic cell separation

Blood was taken from *P. chabaudi* AS infected RAG2$^{-/-}$ and wild-type BALB/c mice, as described above. Packed iRBCs were obtained by centrifugation at 2500 x*g* for 5 minutes at 4 ˚C and 500 μl taken. iRBCs were diluted in 10 ml of MACS buffer (comprising 2 mM EDTA and 2% foetal calf serum in PBS) and passed through a 50 ml XS magnetic cell separation column (MACS, Miltenyi Biotechnology), mounted on a SuperMACS™II Separator magnet (Miltenyi Biotechnology).

Before use, the XS column was washed extensively: with 50 ml distilled water applied against gravity from the top of the column and again from the bottom of the column, to remove air bubbles and contaminants, followed by 50 ml MACS buffer, for equilibration. The column was then mounted upon the magnet, and iRBCs slowly allowed to bind the column matrix, by limiting the flow-rate to less than 2 ml/minute. Following this, the column was washed with 100 ml MACS buffer, before the column was removed from the magnet and iRBCs were eluted.

Elution comprised three steps, all applied from a 50 ml syringe: 50 ml MACS buffer, 50 ml air to remove residual liquid, and a final 50 ml MACS buffer to remove outstanding iRBCs. All iRBCs collected from the initial binding and wash steps were applied again over the column, to ensure that the maximal number of iRBCs were purified by this method. In total, 500 µl of packed iRBCs were applied to the column three times. Columns were re-used after extensive washing under gravity using at least 200 ml 0.15 M NaOH, followed by at least 100 ml distilled water, and stored in absolute ethanol. Eluted iRBCs were pelleted by centrifugation at 1200 x*g* for 10 minutes at 4 ˚C. The supernatants were removed and iRBCs were used to make thin blood smears or at -80 °C in TRIZOL reagent (Invitrogen) for subsequent RNA extraction.

3.2.12 Reticulocyte depletion by saponin lysis of RBC

Blood was taken from *P. chabaudi* AS infected RAG2$^{-/-}$ and wild-type BALB/c mice, as described above. Packed iRBCs were obtained by centrifugation at 2500 x *g* for 5 minutes at 4 ˚C and 100 µl taken. iRBCs were washed three times with 900 µl PBS, followed by centrifugation as before. iRBCs were then re-suspended in 900 µl PBS containing 0.015% saponin (Sigma) and incubated, with gentle shaking, for 10 minutes on ice. Un-lysed cells and cellular compartments were present in the pellet fraction after further centrifugation, as before. These were used to make thin blood smears or at -80 °C in TRIZOL reagent (Invitrogen) for subsequent RNA extraction.

Analysis of *globin* transcript depletion after MACS purfication or saponin lysis was assessed by RT-qPCR, all related information is attached in Appendix 3.6.

3.2.13 Removal of leukocytes via cell filtration

Blood was taken from *P. chabaudi* AS infected RAG2$^{-/-}$ and wild-type BALB/c mice, as described above (section 3.2.1). Blood samples were diluted 15-fold in PBS prior to filtration.

CF11 columns were prepared according to the methods of Sriprawat and colleagues (Sriprawat et al., 2009), using 10 ml syringes filled with CF11 cellulose (Whatman Biosciences). Filtration using CF11 and Plasmodipur filters, required that the columns be pre-wetted with at least 3 ml PBS, following which, the diluted blood was applied to

the column and allowed to filter through under gravity. As leukocytes remain trapped within the column matrix, all blood was collected as the eluate. To ensure all iRBC were eluted, a further 20 ml PBS was applied to the columns, under gravity. Eluates were then centrifuged at 1200 x *g* for 5 minutes at 4 ˚C, and iRBC collected for further use.

To compare leukocyte removal by blood filtration using CF11 cellulose or commercially available Plasmodipur filters (Euro-Diagnostica), the total blood volume of three representative *P. chabaudi* AS infected BALB/c mice were filtered using each method. The proportion of cell types in the blood samples before and after filtration were quantified using the Vetscan II platform (Abaxis).

3.3 Results

3.3.1 Measurement of *cir* transcription by RT qPCR

For this study, primer pairs were designed to amplify *cir* genes within cDNA synthesized from *P. chabaudi* RNA, based upon the *cir* annotation described in chapter 2.3.1. In order to design primers that would specifically target the different annotated *cir* sub-families A, B, C, D, E, F and U, over two hundred potential primer sequences were investigated bio-informatically, before 36 appropriate primers were synthesized (Appendix 3.1).

Primer efficacy was first screened using end-point PCR, to determine which primers were able to amplify a product of the expected size. Of the 36 primer pairs, 15 successfully amplified products (Appendix 3.2). 14 of the PCR positive primers were designed using the same strategy: to amplify a region within *cir* exon 2, hence these were comparable and further validation was carried out for these 14 primers.

The sensitivity of primer pairs was then tested via qPCR using gDNA template titration to ensure that no dimers or other artifacts were observed that could hinder the measurement of *cir* gene expression. Template titration also allowed determination of amplification efficiencies for each primer pair during qPCR, shown in Table 10. Unfortunately, analysis of the dissociation curve for each of the primer sets revealed strong evidence for primer-dimer formation in primers D2 and U3. This would likely interfere with quantification of *cir* transcription so these primers were excluded at this point and are not shown in Table 10.

Table 10: Determination of *cir* primer efficiency by quantitative PCR using *P. chabaudi* gDNA titration and linear regression analysis.

Primer pair:	A1	A2	B1	B2	C1	C2	D1	E1	E2	U1	U2	U4
Slope	-3.85	-4.30	-3.07	-3.87	-4.05	-4.28	-4.46	-2.14	-4.67	-4.36	-7.01	-3.71
r^2	0.93	0.87	0.67	0.92	0.85	0.93	0.95	0.33	0.81	0.95	0.73	0.89
E	1.82	1.71	2.12	1.81	1.77	1.71	1.68	2.93	1.64	1.70	1.39	1.86

Slope refers to the gradient of the regression line, whilst E refers to PCR efficiency. Primers are listed in Appendix 3.1

The slope of the line given in Table 10 reflects the rate of product amplification at the exponential phase of the qPCR reaction. A slope of -3.3 would indicate that the template concentration doubled at each cycle, which translates to an efficiency of 2. All

of the *cir* primers had different gradients for the linear regression lines, indicating that they were not equally efficient.

The specificity of designed primers was next determined, to be sure that the primers targeted the expected *cir* gene groups. This was achieved first via BLAST searches of the *P. chabaudi* AS genome during primer design, and then verified experimentally. Experimental validation involved cloning PCR-amplified products into competent *E. coli* and extraction of plasmid DNA by minipreps (Qiagen). Plasmid inserts were sequenced from the M13F primer site within the vector, and mapped via BLAST. The primer specificities determined by both methods are shown in Table 11.

Table 11: Experimentally determined specificities of *cir* primers.

Primer pair	Bio-informatic investigation:	Experimental investigation:	
	Sequences identified by BLAST*	Number of clones sequenced	Sub-family specificity of sequenced clones
A1	Pc_000270, Pc_000030, Pc_000120	2	A (Pc_000270) & B (Pc_001110)
A2	Pc_000270, Pc_000030, Pc_000120	5	A (Pc_000270) & B (Pc_001110)
B1	Pc_040020, Pc_000400, Pc_000080, Pc_001070	0	Too few clones sequenced
B2	Pc_040020, Pc_000400, Pc_000080, Pc_001070	1	Too few clones sequenced
C1	Pc_000140, Pc_114750	5	C
C2	Pc_030030, Pc_070090, Pc_070160	3	Too few clones sequenced
D1	Pc_001130, Pc_120070	3	D (Pc_001130) & A (Pc_000090)
D2	Pc_100040, Pc_120050	4	D (Pc_100040) & A (Pc_000090)
E1	Pc_120060, Pc_104230**,** Pc_000100	3	E (Pc_000100)
E2	Pc_000100, Pc_000420	0	Too few clones sequenced
U1	Pc_040050, Pc_001050	4	U (Pc_040050)
U2	Pc_042030, Pc_060110	3	Too few clones sequenced
U3	Pc_042030, Pc_030210	2	Too few clones sequenced
U4	Pc_040110, Pc_060140, Pc_042030	3	U (Pc_040110)

Primer sequences are given in Appendix 3.1.* BLAST searches carried out in November 2010. The species identifier PCHAS was abbreviated to Pc. The primer specificity was not certain where fewer than four clones were sequenced. The predominant *cir* gene detected experimentally within each sub-family for each primer pair is shown in brackets, although the majority of BLAST matches could be detected experimentally.

Given the high sequence similarity between *cir* genes, described in chapter 2.3.2, none of the primers were completely specific, but recognized 2 - 4 *cir* genes, Table 11. This means that any transcriptional differences detected with these primers may include changes in a number of closely related genes. In total four primer pairs were used for

investigation of *cir* transcription, which were sub-family specific, did not produce primer-dimers and had similar amplification efficiencies during qPCR (except for primer pair E1, which was less efficient than the other *cir* primers, but was sub-family specific). A summary of their qualities is shown in Table 12.

Table 12: Summary of primers used to measure *cir* transcription

Primer pair	Tm of primers	Experimentally determined specificity	Number of clones sequenced	Predicted number of folds in target RNA sequence [+]	PCR efficiency
C1	64.4 ºC 63.0 ºC	PCHAS_000140 PCHAS_114750	5	8 (Using PCHAS_000140)	1.77
E1	63.5 ºC 62.0 ºC	PCHAS_120060 PCHAS_104230 PCHAS_000100	3	7 (Using PCHAS_000100)	1.64
U1	64.2 ºC 64.0 ºC	PCHAS_040050 PCHAS_001050	4	10 (Using PCHAS_040050)	1.70
U4	63.9 ºC 59.9 ºC	PCHAS_040110 PCHAS_060140 PCHAS_042030	3	6 (Using PCHAS_040110)	1.86

The names hereafter used for these primers are given in brackets underneath their number. Tm refers to the melting temperature for each primer. [+] Number of predicted RNA folds in amplified target sequence predicted by Mfold (http://frontend.bioinfo.rpi.edu/applications/mfold/cgi-bin/rna-form1-2.3.cgi) using default settings. Primer sequences are listed in Appendix 3.1.

The feasibility of measuring *cir* gene transcription using the four primer pairs: U1, U4, C1 and E1 described above (Table 12), was investigated in three experiments shown in Figure 21 and Figure 22. Immuno-competent (BALB/c) and immuno-deficient ($RAG2^{-/-}$) mice were infected intra-peritoneally with $1x10^5$ *P. chabaudi* iRBCs. Blood was harvested immediately prior to schizogony (11.30 - 12.00 in reverse light conditions) at the peak of infection, on day 7 after infection. As these were pilot experiments, only three mice per group were analyzed. RT-qPCR reactions having low confidence values were excluded from these figures (marked with x, and tabulated in Appendix 3.4), which in conjunction with the small group size, prevented statistical analyses since no changes were observed in all mice.

Levels of *cir* transcripts were compared relative to the mouse in which parasites were expanded (the donor), normalized to *beta tubulin*. Changes in *cir* transcript levels were expressed as a ratio between the donor and recipient mice. Ratios approximately equal to one indicated no change or only very small changes in *cir* transcription, thus only ratios with more than two-fold change were considered. A ratio of 2 would indicate a two-fold increase in *cir* transcripts (1:2), in the recipient compared to the donor *P. chabaudi* parasite population. Conversely, a ratio of 0.5 would indicate a two-fold decrease (2:1),

in the recipient compared to the donor. For ease of view, these data are represented using a $\log_2$ scale, where a two-fold change in either direction is displayed at an equal scale.

Figure 21a shows *cir* transcription in each of three BALB/c mice, which were infected from a RAG2$^{-/-}$ donor mouse. The *P. chabaudi* populations present in all BALB/c mice appeared to display minor (less than two-fold) changes compared to the donor RAG2$^{-/-}$ mouse. The only substantial differences were observed in one or more of the genes targeted by the U4 primers: PCHAS_040110, PCHAS_060140 and PCHAS_042030. These primers detected 2-12 -fold lower *cir* levels in all three *P. chabaudi* infected BALB/c mice, compared to the donor RAG2$^{-/-}$ mouse.

In a second experiment, three RAG2$^{-/-}$ mice were infected with parasites from a BALB/c mouse (mouse 1 from the previous experiment); Figure 21b. This experiment aimed to determine whether *cir* expression changed during *P. chabaudi* infection in the absence of lymphocytic selection. As previously observed, few differences were detected in comparison to the donor mouse. The only substantial changes were observed in the genes targeted by U1 primers: PCHAS_040050 and PCHAS_001050, which detected almost four-fold higher *cir* levels in two of the three RAG2$^{-/-}$ mice.

To determine whether *cir* transcription changed during a single 24-hour intra-erythrocytic growth cycle, RT-qPCR was carried out using blood samples taken at four points during the cycle, Figure 22. The time-points contained parasites that were predominantly ring or trophozoite stages (defined in Table 13) examples of each stage are shown in Figure 22a. Mature *P. chabaudi* schizonts are rarely found in the periphery because they sequester (Cox et al., 1987, Gilks et al., 1990).

Table 13: Parasite stages observed during intra-erythrocytic development.

Time-point	Stage composition of *P. chabaudi* iRBC (%):								
	Ring stage			Trophozoite stage			Schizont stage		
Mouse	**1**	**2**	**3**	**1**	**2**	**3**	**1**	**2**	**3**
Early trophozoite 8.30 h	25.75	18.23	22.50	74.25	81.77	77.50	0.00	0.00	0.00
Mid trophozoite 10.30 h	12.04	12.50	13.59	87.04	86.67	85.87	0.93	0.83	0.54
Late trophozoite 12.30 h	13.82	20.86	10.52	83.87	76.26	86.91	2.30	2.88	2.58
Ring 22.30 h	89.72	58.82	91.54	9.35	41.18	6.34	0.93	0.00	2.11

Blood samples were taken at the indicated time-points on day 7 of a *P. chabaudi* infection of three BALB/c mice, kept in reverse light conditions (where schizogony occurred between 11.30 and 12.30 daily). Each time-point corresponded to a particular stage of parasite development, for the majority of iRBCs. Counts were performed by microscopy, enumerating the parasites observed within 10,000 RBCs.

Here, *cir* transcription was compared within the same mouse between the time-points representing predominantly: mid trophozoites, late trophozoites and ring stages, relative to the first time-point, early trophozoites; all normalized to *beta tubulin*. It is likely that *cir* genes were least transcribed in early trophozite stage parasites, as the trend observed in Figure 22 was for increased levels of *cir* transcripts in other stages of parasite development, compared to this time-point. However, further transcriptional patterns could not be identified for any particular time-point or sub-set of *cir* genes, indicating that *cir* transcription is a dynamic process *in vivo* during intra-erythrocytic parasite development.

3.3.2 Measurement of *cir* transcripts by RNA sequencing

A major limitation for the use of RT-qPCR analysis to measure *cir* transcript levels during *P. chabaudi* infection was that very few of the *cir* family members could be effectively analyzed (as observed in section 3.3.1). Recent advances in high throughput sequencing technologies have enabled their application to transcriptome level analyses, and may outcompete the use of microarrays for this purpose (Shendure, 2008). To determine whether Solexa / Illumina whole transcriptome sequencing could be applied to *P. chabaudi* for the measurement of *cir* transcription, two trial RNA samples were prepared. The RNA was extracted directly from the whole blood volume of two BALB/c mice, taken at the peak of separate *P. chabaudi* infections. The results are summarized in Table 14.

Table 14: Statistics for trial *P. chabaudi* RNA sequencing samples.

	BALB/c sample 1	**BALB/c sample 2**
Parasitaemia	42.93	40.78
Sample composition:		
Mus musculus *	23.23%	81.33%
rRNA	3.10%	0.01%
P. chabaudi mRNA	73.67%	18.66
***P. chabaudi* statistics:**		
Total number of sequence reads	3038218	2825036
Mapped sequence reads	20%	34%
Reads with 40 bp overlap	269111	169499
Number of perfectly mapping reads ^	163363	18771
Reads which contained repetitive sequence	6%	21%
Unique reads, matching a single gene	14%	18%
Total bp mapped to *P. chabaudi* AS genes	16612855	15923935
Proportion of genes not covered by any reads +	89%	99%

The parasitaemia of each mouse is shown on the day blood was taken and RNA extracted. Read lengths of 54 nucleotides were generated. * Including DNA and all RNA species. ^ Containing no mis-matches with the genome sequence. + The total length of mapped transcripts as a percentage of all annotated genes (covering 18832209 base pairs).

Both RNA samples contained high levels of host (mouse) RNA: 23.23 - 81.33% of the total sample (Table 14). One of the samples also contained a substantial percentage of ribosomal RNA (rRNA) contamination (3.10%, Table 14). These sequences were excluded bio-informatically and the remaining sequences were used to determine *P. chabaudi* AS specific statistics for the two samples. Between 1 and 10 million *P. chabaudi* sequence reads were obtained, with a read length of 54 base pairs (bp). After mapping these reads to the *P. chabaudi* genome sequence, the majority of the genome

was not covered by a single sequence read (89-99%; Table 14, the list of *cir* genes detected is attached in full in Appendix 3.5). Thus, it is likely that not all *P. chabaudi* transcripts were detected from these samples, and the levels of each transcript could not be determined with confidence.

Nonetheless, Figure 23 shows that approximately two thirds of *cir* genes could be detected from these two RNA samples. Dominant *cir* genes were expressed at between four and ten times the level of the next most highly transcribed *cir*. Whole transcriptome sequencing data from the first sample were obtained during the second stage of *cir* gene annotation (described in chapter 2.2.1). Sequencing reads for 73 *cir* genes were detected, corresponding to approximately 62% of the *cir* repertoire. Subsequently, data were obtained for the second sample after *cir* gene annotation was complete, thus more *cir* genes could be detected in this sample: 122 of the total 199 *cir* genes. Notably, in both samples one transcript dominated the expressed *cir*s: PCHAS_010120 in sample 1 (excluded from CIR phylogenetic analyses), and PCHAS_110030 in sample 2 (CIR sub-family U).

The detection of a dominant *cir* transcript in both *P. chabaudi* RNA samples could be due either to the majority of iRBCs transcribing that *cir* gene at low levels or a few iRBCs transcribing that *cir* at very high levels. Other highly expressed *cir* genes (with greater than 200 sequence reads detected) were also detected in the second sample: PCHAS_070130, PCHAS_083690 and PCHAS_040060. All highly expressed *cir* genes are indicated with stars in Figure 23.

The ability to detect many *cir* transcripts, including a dominant *cir*, by RNA sequencing of *P. chabaudi* infected blood was encouraging. However, there was considerable room for improvement. The proportion of the genome covered by sequence reads was low, even to 1x coverage, Table 14. Transcriptome coverage to at least 4x redundancy is required in order to confidently identify transcripts and other phenomena that may be detected by RNA sequencing, such as novel splice variants (Trapnell et al., 2009). An RNA sequencing experiment is able to detect a finite number of reads, so the presence of high abundance mouse and rRNAs would have reduced the capacity for detection of *P. chabaudi* transcripts. Therefore, the further use of RNA sequencing technology required removal of high abundance contaminating RNA species, which precluded the detection of *P. chabaudi* transcripts.

Efforts were first made to investigate the source of the *Mus musculus* contaminating RNA, which comprised both rRNA and mRNA species. In fact, the host contamination was visible by agarose gel electrophoresis of the *P. chabaudi* RNA samples (prior to rRNA digestion), shown in Figure 24. Whilst mRNA and several rRNA species from *P. chabaudi* and *M. musculus* were indistinguishable on the gel, one rRNA species was detected solely in *P. chabaudi* infected samples (labelled band 2 in Figure 24). By contrast, band 1 was present at higher concentration in RNA samples from naïve mice than *P. chabaudi* infected mice, compared to the other rRNA bands (Figure 24a and b). Thus, a high degree of host contamination in *P. chabaudi* RNA samples could be visualized after gel electrophoresis by the presence of rRNA band 1. Lymphocytes were unlikely to be the source of this host contamination, as bands of equal intensity were observed in both BALB/c and $RAG2^{-/-}$ mice, Figure 24b. The source of host contamination was also not induced by *P. chabaudi* AS infection, as host RNA could be clearly seen in naïve mice, Figure 24c.

To deplete leukocytes present in *P. chabaudi* infected blood, filtration of blood samples was carried out using CF11 columns and commercially available Plasmodipur filters (EuroProxima), Table 15.

Table 15: Leukocyte depletion by filtration of *P. chabaudi* infected blood.

Total leukocyte counts ($x10^7$):						
	Plasmodipur			**CF11**		
Mouse	**1**	**2**	**3**	**1**	**2**	**3**
Before filtration	4.29	4.56	3.00	4.28	4.35	6.03
After filtration	0.04	0.12	0.05	0.47	0.48	1.22
Total loss (%)	99.07	97.37	98.50	89.02	88.97	79.77
Total RBC counts ($x10^9$):						
	Plasmodipur			**CF11**		
Mouse	**1**	**2**	**3**	**1**	**2**	**3**
Before filtration	7.71	5.58	6.00	6.50	6.08	5.95
After filtration	7.66	5.40	5.64	4.98	4.43	5.03
Total recovery (%)	99.35	96.77	94.00	76.62	65.15	84.54

The proportion of leukocytes and RBCs present in each sample of *P. chabaudi* infected blood, from three BALB/c mice at day 7 of infection, was enumerated using the VetScan II (Abaxis), before and after filtration using either a Plasmodipur filter (EuroProxima) or a CF11 column.

Plasmodipur filtration consistently removed a higher proportion of leukocytes from *P. chabaudi* infected blood samples (12.39% more than after CF11 filtration, Table 15),

with a greater recovery of RBCs (21.27% more than after CF11 filtration, Table 15). However, after more than 97% leukocyte depletion by Plasmodipur filtration, host RNA contamination was still visible upon gel electrophoresis of *P. chabaudi* RNA samples, Figure 24d. Thus, it was concluded that leukocytes were not the source of the host RNA contamination.

The contribution of RNA from red blood cells was then investigated as a potential source of contamination for *P. chabaudi* RNA samples. High numbers of reticulocytes could be observed in the blood of both uninfected and *P. chabaudi* AS infected mice, Table 16. Reticulocytes in the blood of BALB/c and RAG2$^{-/-}$ mice were detected using two methods. First, flow cytometry was performed using the same blood samples, using the stain DRAQ5 (Biostatus Limited), which only detects DNA, and thiazole orange (Sigma-Aldrich), which detects all nucleic acids; Figure 25a. In addition, thin blood films were counter stained with Cresyl brilliant blue (BDH) to detect the reticular material and Giemsa (VWR international) to detect nuclei, Figure 25b. Two hundred RBCs were counted, and the proportion of reticulocytes recorded. In both strains of mice the reticulocyte numbers were similar (with a mean of 0.84% in naïve and 1.94% in *P. chabaudi* infected mice), within the acceptable range for healthy mice, up to 3% (McGarry et al., 2010).

Table 16: The proportion of reticulocytes in the blood of naïve and *P. chabaudi* infected mice.

Mouse strain (Age, weeks)	**% Reticulocytes (enumerated by:)**						**Parasitaemia**		
	Cresyl blue			**Flow cytometry**					
Mouse	**1**	**2**	**3**	**1**	**2**	**3**	**1**	**2**	**3**
BALB/c (9)	0.67	0.34	0.83	0.76	0.72	0.77	Naïve	Naïve	Naïve
BALB/c (5)	1.19	3.00	1.00	2.76	3.91	2.88	19%	19%	17%
RAG2$^{-/-}$ (5)	2.00	0.60	1.10	1.00	0.54	0.72	Naïve	Naïve	Naïve
RAG2$^{-/-}$ (10)	3.50	1.50	1.33	5.32	1.19	1.08	25%	< 1%	2%

Tabulated data showing the percentage of RBC that were reticulocytes in different strains of naïve and *P. chabaudi* infected mice, counted using either microscopy or flow cytometry. Three individuals were analyzed per group, samples taken on day 7 of infection. The parasitaemia of infected mice is also given.

Table 16 shows that reticulocyte numbers were increased in *P. chabaudi* infected mice compared to naïve mice, and that RAG2$^{-/-}$ mice appeared to have slightly higher numbers of reticulocytes than BALB/c mice. However, a large degree of variation between mice made conclusions difficult to draw. This may have been due to different

definitions of 'a reticulocyte' by the two methods, whereby DNA positive *P. chabaudi* infected reticulocytes were excluded by flow cytometry analysis, but could be detected by microscopy. Equally, the proportion of reticular material detected by each method may have been different. To minimize variation, mice aged between 6-8 weeks were used in all subsequent experiments.

To measure the amount of host mRNA contamination in *P. chabaudi* RNA samples, all transcripts identified from the initial RNA sequencing samples were examined (data not shown). The majority of these host contaminants were *globin* transcripts, in agreement with the high percentages of reticulocytes observed in the blood of *P. chabaudi* infected mice. Primers were designed to amplify *Mus musculis beta globin*, to determine the extent of *P. chabaudi* RNA contamination by RT-qPCR.

Beta globin primers were validated as described for *cir* primers (section 2.2.5, Appendix 3.6). Their specificity was assessed by cloning of PCR-amplified products; all three clones analyzed contained the *beta globin* gene. RT-qPCR was then used for determination of primer efficiency. As these primers spanned the first exon: exon boundary of the *beta globin* gene, PCR efficiency was calculated by titration of a pooled cDNA standard rather than using gDNA. This indicated that the *beta globin* primers were ten fold less efficient than those for the *P. chabaudi* gene *beta tubulin*: having PCR efficiencies of 20.09, and 2.00, respectively (calculated using the equation given in section 3.2.8.i). However, different levels of these transcripts may have been present in the cDNA pool used for estimation of primer efficiency and for this reason, correction for PCR efficiency was not applied in the estimation of their transcript levels. This meant that the relative levels of *beta globin* measured by RT-qPCR could have been actually ten-fold higher than the *beta tubulin* transcript levels in the analyses described below. Thus, any detection of *beta globin* transcripts by RT-qPCR would indicate substantial host contamination of *P. chabaudi* mRNA.

Different methods were carried out in an attempt to reduce reticulocytes from infected blood. The first approach was to use a high dose infection as it is known that *P. chabaudi* infection with the standard intra-peritoneal (i/p) inoculum of $1x10^5$ iRBC induces an efflux of reticulocytes from the bone marrow, known as reticulocytosis, around the peak of infection (Podoba and Stevenson, 1991). Reticulocyte numbers were compared between mice given the standard inoculum ($1x10^5$ iRBCs) and mice given a high dose inoculum of $1x10^8$ iRBCs, with the idea that sufficient parasite material might be collected before induction of reticulocytosis. The percentage of reticulocytes in the

blood of these mice was counted by both cresyl blue microscopy and flow cytometry, Table 17. However, no differences were observed in the proportion of reticulocytes for the two groups.

RT qPCR was used to measure the level of mouse *beta globin* contamination in *P. chabaudi* RNA samples, as a proportion of the constitutively expressed *P. chabaudi* gene *beta tubulin* (used to represent the level of *P. chabaudi* transcripts present). Figure 25c shows that host contamination of parasite material was reduced in mice infected with $1x10^8$ compared to $1x10^5$ iRBCs, however there was still more *beta globin* than *beta tubulin* RNA present in the sample. Given that such a high dose infection was not practical for most *P. chabaudi* experiments, and that the early rise of parasitaemia did not sufficiently reduce the *globin* to parasite RNA ratio, this avenue was not pursued in further experiments.

Table 17: The proportion of reticulocytes in the blood of *P. chabaudi* infected mice.

Strategy	% Reticulocytes (enumerated by:)						Parasitaemia		
	Cresyl blue			Flow cytometry					
Mouse	**1**	**2**	**3**	**1**	**2**	**3**	**1**	**2**	**3**
10^5 infectious dose	0.80	0.41	0.85	0.93	0.75	0.89	12.0%	10.4%	12.8%
10^8 infectious dose	1.70	0.92	4.55	1.33	0.74	0.92	27.6%	28.8%	37.6%

Tabulated data showing the percentage of RBC that were reticulocytes in BALB/c mice infected with different doses of *P. chabaudi* iRBCs. Three individuals were analyzed per group, samples were taken on day 7 of infection for mice given the 10^5 infectious dose, and day 4 of infection for mice given the 10^8 infectious dose. Reticulocytes were counted using either microscopy or flow cytometry.

Several investigators have made use of the magnetic property of haemozoin to enrich *Plasmodium* iRBCs and deplete other cell types [described by (Paul et al., 1981)]. This approach was explored for *P. chabaudi* using a large MACS column in combination with the SuperMACSII magnet (Miltenyi Biotechnology). Whilst the proportion of reticulocytes in blood was lower in the MACS purified blood samples, there was a substantial loss of material, Table 18. A volume of 100 µl packed RBCs was purified over the column twice, for three individual mice, yet only 3-10 µl of packed RBCs were recovered. Not only were the purified cells less uniform than is usual for Giemsa-stained iRBC, but despite extensive washing of the column, uninfected cells, including reticulocytes, were still present in the purified iRBC fraction (48 - 64% of cells were RBCs, and 0.77 – 5.88% of these were reticulocytes). An example of purified cell morphology is shown in Figure 25d. The proportion of *beta globin* relative to *beta*

tubulin was again measured by RT-qPCR, but there was no difference between the samples before and after MACS purification, Figure 25d.

Enrichment of the intra-erythrocytic parasites was attempted using saponin lysis of erythrocyte membranes. This process forms pores in both the erythrocyte and PV membranes of *Plasmodium* iRBC, as they are derived from the same source (Pouvelle et al., 1994, Ward et al., 1993), but leaves the developing parasites and their nuclei intact (Christophers and Fulton, 1939). Using this method, the proportion of reticulocytes decreased substantially relative to the number of parasites counted in two samples. However, in the other sample there was virtually no difference (denoted * in Table 18), indicating that the lysis process was highly variable for *P. chabaudi* iRBCs. Similarly to MACS purification, saponin mediated cell lysis resulted in approximately ten-fold loss of material, Table 18, yet no difference was observed in the proportion of *beta globin* relative to *beta tubulin* found in the extracted RNA, Figure 25d. This signified that neither method was able to reduce the proportion of *M. musculus* contamination in the *P. chabaudi* RNA samples.

Table 18: The proportion of reticulocytes in *P. chabaudi* infected blood samples, subjected to purification by MACS or saponin lysis.

Before separation:			After MACS separation:		After saponin lysis:	
Reticulocytes	**Parasitaemia**	**P:R ratio**	**Volume of eluted iRBC**	**P:R ratio**	**Volume of eluted iRBC**	**P:R ratio**
1.50 %	16.5%	70.92	10 ul	11.00	8 ul	100.00
1.50 %	12.5%	18.00	6 ul	8.33	16 ul	9.52 *
2.00 %	14.0%	24.01	3 ul	7.00	12 ul	66.67

The percentage of RBC that were reticulocytes and iRBC, before and after treatment, were counted by microscopy. As many uninfected cells were lost after treatment, the number of reticulocytes was counted within fields containing 200 parasites, and this was divided by the number of reticulocytes counted to give a ratio of parasites per reticulocyte (P:R). Three BALB/c mice were analyzed per group, samples taken on day 7 of infection. 100 µl packed RBC were used for all experiments and the volume of packed RBC recovered after treatment was noted as the volume of eluted iRBC. *Poor lysis, discussed above.

In conclusion, although a major source of *P. chabaudi* mRNA contamination was derived from host cells, specifically reticulocytes, attempts to deplete these cells from infected blood samples were unsuccessful. The identification of improved methods, both to remove host mRNA contamination, and to reduce the levels of the other major contaminant (rRNA) from *P. chabaudi* RNA samples, will be required in order to carry out future RNA sequencing experiments.

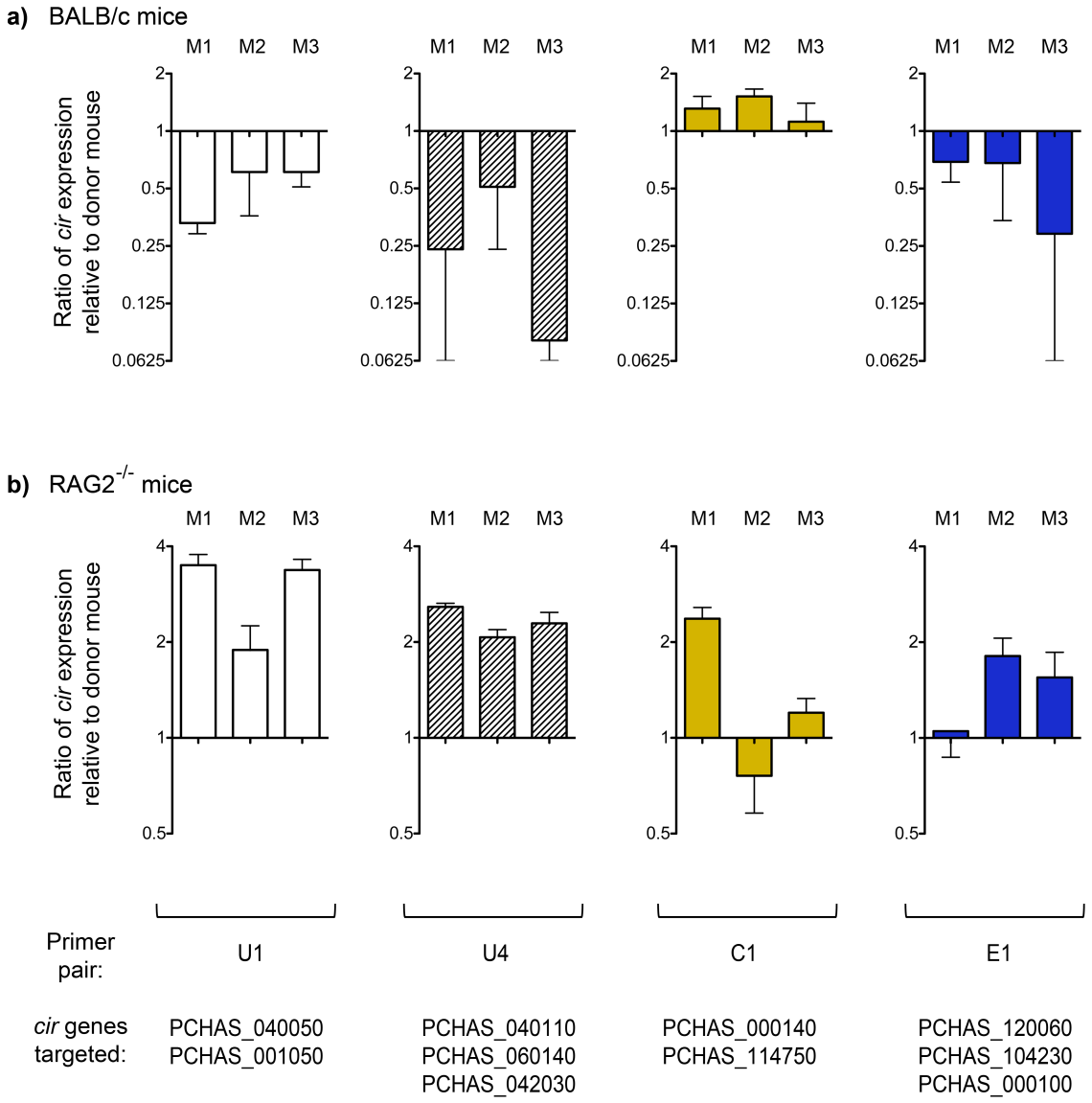


Figure 21: RT qPCR analysis of *cir* transcription during *P. chabaudi* AS infection.

a) *P. chabaudi* infection of three BALB/c mice (labelled M1, M2, M3), infected with parasites expanded in a RAG2$^{-/-}$ donor mouse.

b) *P. chabaudi* infection of three RAG2$^{-/-}$ mice (labelled M1, M2, M3), infected with parasites expanded in a BALB/c donor mouse.

5 ng RNA was analyzed from *P. chabaudi* infected blood, taken on Day 7 of infection. The ratio of *cir* transcription normalized to the parasite reference gene, *beta tubulin*, is shown for each mouse, relative to the *P. chabaudi* donor mouse. For ease of view, these data are represented using a $\log_2$ scale, where a two-fold change in either direction is displayed at an equal scale. Ratios are not displayed as $\log_2$ values however.

Primers used were: U1, U4, C1 and E1, given under the respective graphs, beneath which the *cir* genes amplified by each primer pair are listed.

In all graphs the mean of three technical replicates is plotted, data are tabulated in Appendix 3.4. Error bars represent standard deviation.

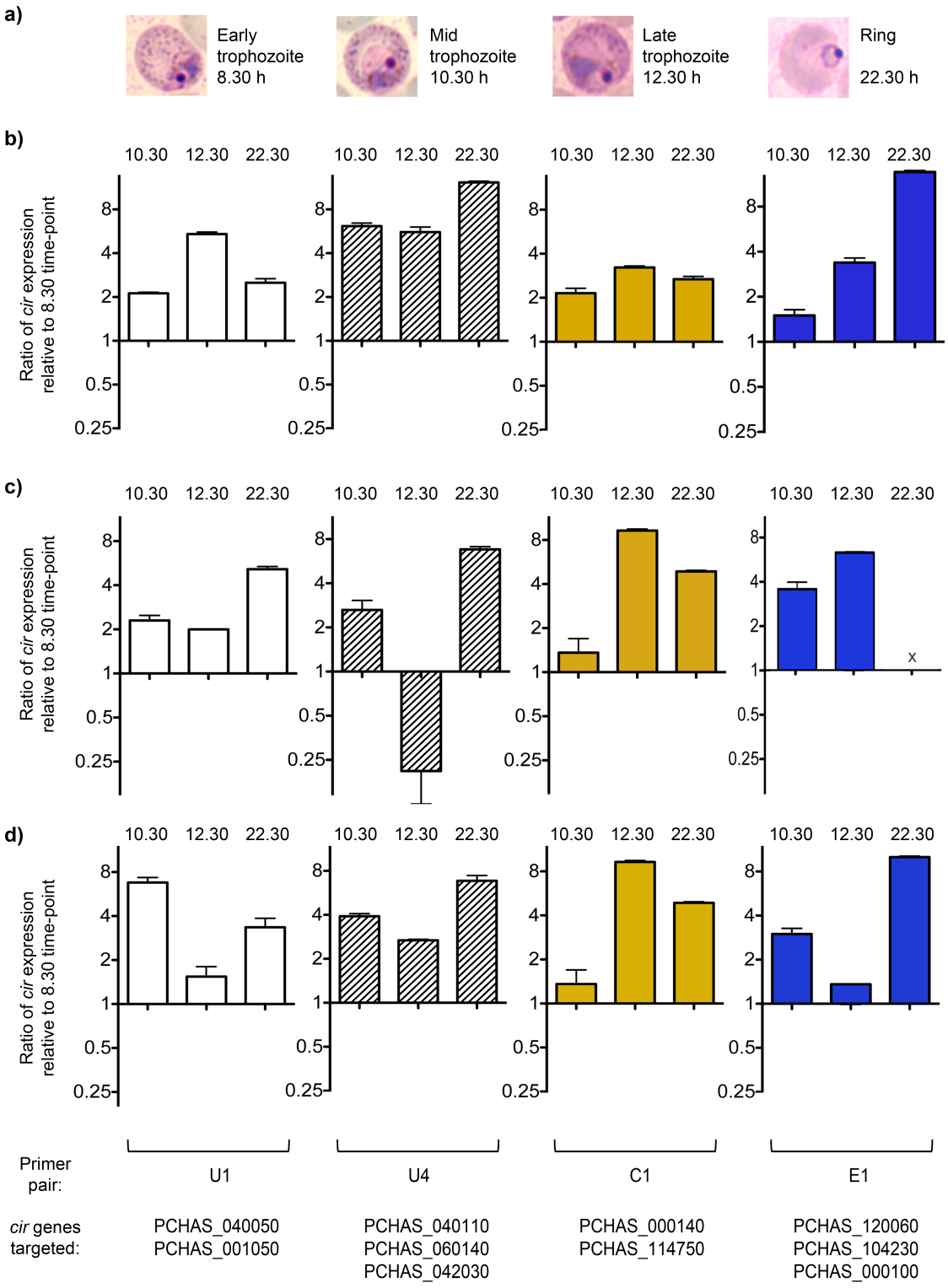

a)
Early trophozoite 8.30 h
Mid trophozoite 10.30 h
Late trophozoite 12.30 h
Ring 22.30 h
b)
c)
d)
10.30 12.30 22.30
Ratio of cir expression relative to 8.30 time-point
8
4
2
1
0.5
0.25
x
Primer pair:
U1
U4
C1
E1
cir genes targeted:
PCHAS_040050
PCHAS_001050
PCHAS_040110
PCHAS_060140
PCHAS_042030
PCHAS_000140
PCHAS_114750
PCHAS_120060
PCHAS_104230
PCHAS_000100

Figure 22: RT qPCR analysis of *cir* transcription in different stages of *P. chabaudi*.

a) A representitive *P. chabaudi* iRBC is shown for the following stages of intra-erythrocytic development: early trophozoite (8.30), mid trophozoite (10.30), late trophozoite (12.30), and ring (22.30), as described in Table 13.

5 ng RNA was analyzed from three *P. chabaudi* infected BALB/c mice, on Day 7 of infection. *cir* transcription was measured in each mouse, shown in **b)**, **c)** and **d)**.

The ratio (arbitrary units) of *cir* transcription is shown for the stages: early trophozoite, late trophozoite and early ring, relative to *cir* transcription in the same mouse at the late ring stage of iRBC development. All *cir* transcription was normalized to the parasite reference gene, *beta tubulin*. For ease of view, these data are represented using a $\log_2$ scale, where a two-fold change in either direction is displayed at an equal scale. Ratios are not displayed as $\log_2$ values however.

In all graphs the mean of three technical replicates is plotted. Error bars represent standard deviation. Mice were excluded where the standard deviation between technical replicates exceeded 0.775 (80% confidence), indicated by x. All data are tabulated in Appendix 3.4.

Primers used were: U1 (white bars), U4 (hashed bars), C1 (yellow bars) and E1 (blue bars). These are given under the respective graphs, beneath which the *cir* genes amplified by each primer pair are listed.

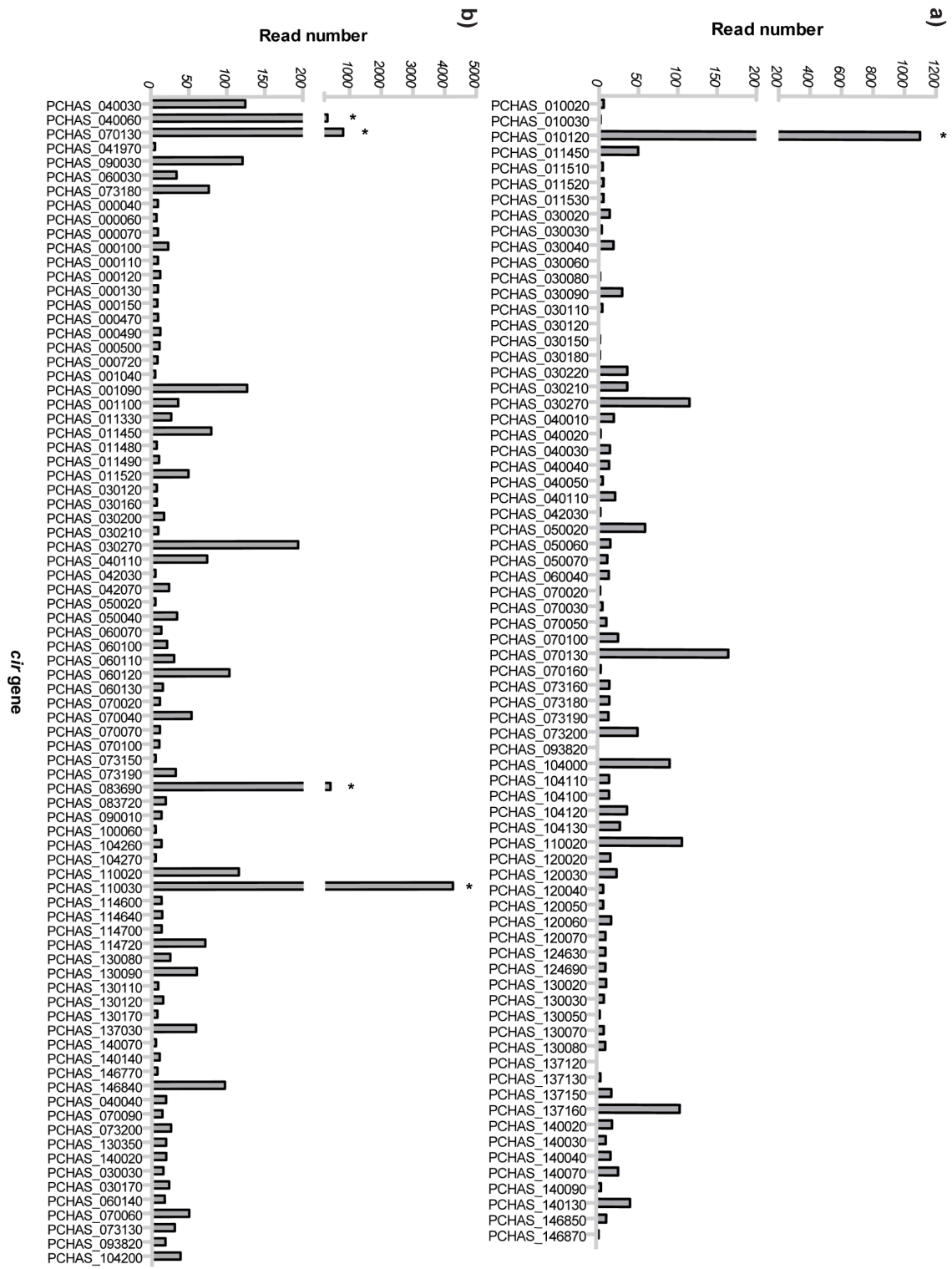


Figure 23: Detection of *cir* transcription by Solexa/ Illumina RNA sequencing.

The total number of sequence reads detected for each *cir* gene is plotted.

a) Sample 1 shows all *cir* genes for which sequence reads were detected.

b) Sample 2 shows *cir* genes for which more than 5 sequence reads were detected.

Highly expressed *cir* genes, where at least 200 transcripts were detected, are indicated by *. All cir transcripts are tabulated in Appendix 3.5.

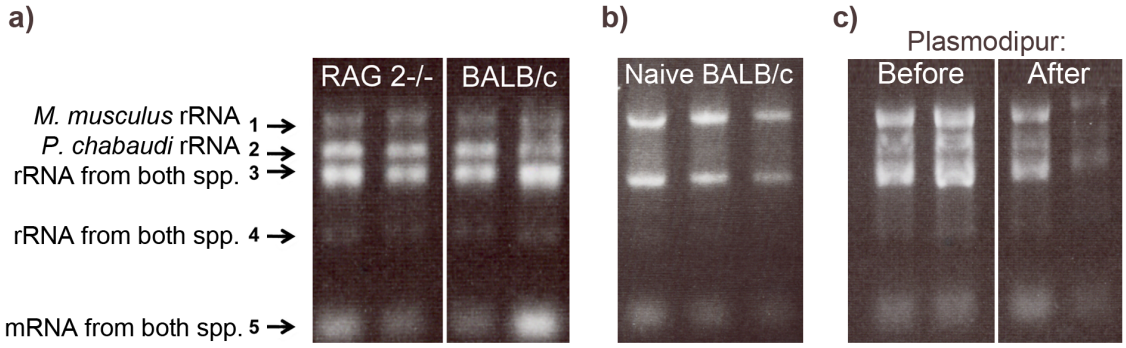


Figure 24: Host contamination in *P. chabaudi* AS RNA despite leukocyte depletion.

a) *P. chabaudi* RNA extracted from infections in two RAG2$^{-/-}$ and BALB/c mice. Arrows represent the different RNA species resolved on this gel: 1), 3) and 4) *Mus musculus* ribosomal (r) RNA, 2), 3) and 4) *P. chabaudi* rRNA, 5) mRNA. RNA from both species was observed in bands 3), 4) and 5), where the word species was abbreviated to spp.

b) RNA extracted from three uninfected BALB/c mice.

c) *P. chabaudi* RNA extracted from infections of two BALB/c mice before and after leukocyte depletion by Plasmodipur filtration (the efficiency of which was assessed by microscopy, where no leukocytes were counted in more than 25,000 RBCs).

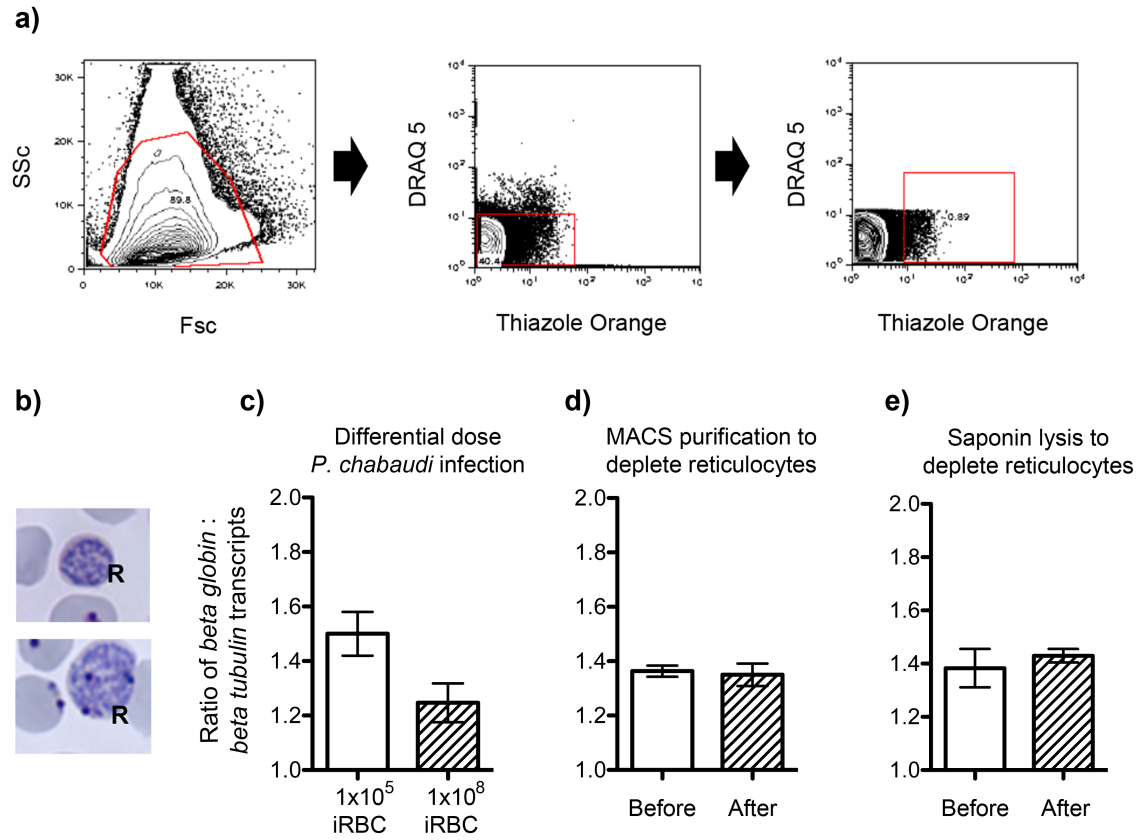


Figure 25: Strategies for depletion of reticulocytes from *P. chabaudi* infected blood.

Reticulocyte numbers were estimated using two methodologies, outlined below:

a) Example analysis of flow cytometry data. Dead cells were excluded according to their size (forward scatter, FSc) and granularity (side scatter, SSc) as defined by the red shape, shown in the left panel. DNA positive cells were excluded on the basis of DRAQ5 staining (a DNA specific dye), as defined by the red reactangle, centre panel. Remaining RNA positive cells were enumerated according to Thiazole Orange staining (staining both RNA and DNA), as defined by the red rectangle, right panel.

b) Examples of RBCs from *P. chabaudi* infected blood, stained with Giemsa and counter-stained with cresyl blue. Reticulocytes are denoted by R.

The expression levels of *beta globin* (*M. musculus*) and *beta tubulin* (*P. chabaudi*) were determined for the experiments outlined in **c)**, **d)** and **e)** by RT-qPCR. The mean ratios of *beta globin* : *beta tubulin* transcripts present in the blood of *P. chabaudi* infected mice are plotted. Error bars represet the standard error of the mean.

c) Three mice were infected with *P. chabaudi* i/p, with an infectious dose of either 10^5 or 10^8 iRBCs. Blood samples were taken on day 7 and 4 of infection, respectively.

Blood samples taken on day 7 of a standard *P. chabaudi* infection (10^5 infectious dose i/p), were analyzed by RT-qPCR before and after the following treatment: iRBC enrichment by MACS, **d)**, and lysis of uninfected RBCs using saponin, **e)**. Three mice were analyzed per group.

3.4 Discussion

In this chapter, *cir* transcripts were measured in *P. chabaudi* infected blood samples using both RT-qPCR and RNA sequencing. Four pairs of primers were designed that could specifically amplify *cir* genes within the sub-families C, E and U. These primers were used in preliminary RT-qPCR analyses of *cir* transcription during *P. chabaudi* infections of BALB/c and RAG2$^{-/-}$ mice, finding variable levels of *cir* transcripts between mice. In addition, RNA sequencing detected transcription of two thirds of the *cir* repertoire at the peak of *P. chabaudi* infection, despite high levels of host contamination. One *cir* gene appeared to dominate each *P. chabaudi* infection, which was expressed at 4-10 times the level of the next most highly transcribed *cir*. The dominant *cir* may be transcribed by the majority of iRBCs at low levels, or by fewer cells at high levels. In both RNA sequencing experiments, the dominant *cir* belonged to sub-family U, although members of all the *cir* sub-families identified in chapter 2.3.2 were detected in both samples. Members of the sub-families C, E and U were also detected during the blood stages of *P. chabaudi* infection by RT-qPCR. This has also been reported in another rodent model, *P. yoelii*, using group specific primers for RT-qPCR (Fonager et al., 2007).

A dynamic profile of *cir* expression was observed throughout the blood stages of parasite development, although one trend could be observed. The lowest levels of *cir* transcripts were detected in early trophozoite stage iRBCs, as all other stages had comparably higher *cir* detection. This agrees in part with previously published Northern blot data, which detected the highest levels of *cir* transcripts in late trophozoite stages (Janssen et al., 2002). Increased levels of some *cir* transcripts, notably those targeted by the primer pair U4, could also be detected in ring stage parasites, compared to the early trophozoite time-point. This may be explained by increased parasite biomass in this time-point after RBC invasion, compared to the trophozoite stages.

An important consideration for RT-qPCR experiments is the choice of reference gene against which normalization is performed to provide a relative calculation of transcript abundance. In rodent-infecting species of *Plasmodium*, very few primers have been designed and validated for quantitative PCR, giving a limited choice of possible reference genes. Primers targeting *beta tubulin* have been used previously in this regard to normalize transcriptional measurements of *yir* genes (Cunningham et al., 2009). Expression data for *beta tubulin* throughout the erythrocytic growth cycle has been

collected for both *P. falciparum* (Le Roch et al., 2003, Llinás et al., 2006, Otto et al., 2010, Bozdech et al., 2003a) and *P. vivax* (Bozdech et al., 2008), whose results were almost identical. A diagram of the *P. vivax beta tubulin* expression pattern in three patient isolates is shown in Figure 26.

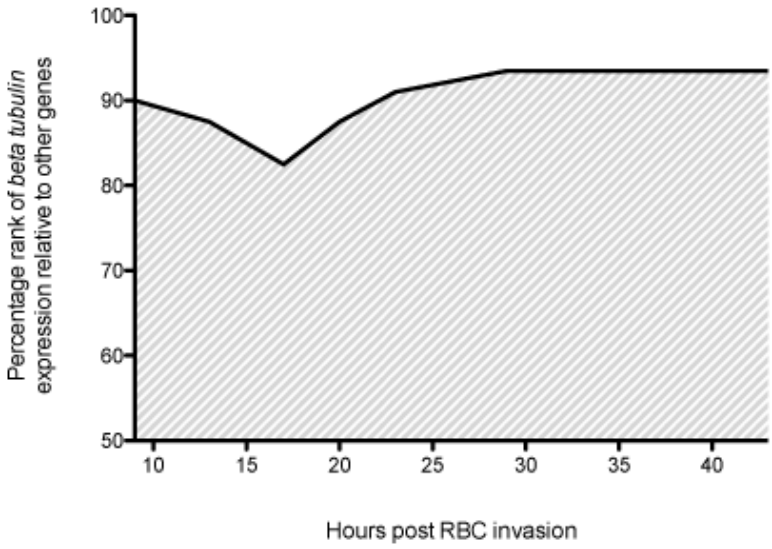


Figure 26: Transcriptional profile throughout the *P. vivax* 48 h intra-erythrocytic development cycle.

Adapted from (Bozdech et al., 2008). Best fit Pearson correlations were used to correlate *P. vivax beta tubulin* expression intensity to the expression data in time-points 9, 13, 17, 20, 23, 29, 35, 40, and 43 h post RBC invasion in the *P. falciparum* transcriptome. This graph shows an approximate average from the values obtained from 3 patient samples (Bozdech et al., 2008).

Whilst *beta tubulin* transcription remained fairly constant during the erythrocytic cycle, Figure 26 (Bozdech et al., 2008), the percentages were slightly lower at the early time-points post RBC invasion (approximately 90 - 95% of the transcription observed for other *P. vivax* genes). A similar pattern of *beta tubulin* transcription is also likely to be observed during *P. chabaudi* intra-erythrocytic development. This may have contributed to the higher levels of *cir* transcripts detected in ring stage iRBCs compared to other stages of development.

Therefore, as the normalization would be slightly different at these early stages it would be more accurate in future experiments to compare *cir* expression relative to *beta tubulin* within the equivalent stage of iRBC development, rather than between stages. The *cir* primers validated here may be used for further experiments where differences in transcription would be expected in particular *cir* genes, or where the quantities of material available would be too low for analysis of the whole transcriptome. For

example, analyses of *cir* transcription could be performed at several time-points during a chronic *P. chabaudi* infection to investigate whether *cir* transcription changes in line with the antigenic properties of the iRBC surface [as described by (McLean et al., 1982, McLean et al., 1986)]. In addition, *cir* transcription could be measured within individual iRBCs. Different sub-types of *rif* and *pir* genes can be co-expressed within individual iRBCs (Fonager et al., 2007, Petter et al., 2007, Bozdech et al., 2008), whilst infections of RAG2$^{-/-}$ mice with individual *P. yoelii* iRBCs have given rise to parasites expressing different *yir* profiles (Cunningham et al., 2009), indicating that the parasites must display plastic *pir* transcription.

Early experiments have shown that *yir* transcription changes substantially during a secondary *P. yoelii* infection of immuno-competent mice, but not in immuno-compromised animals (Cunningham et al., 2005), indicating that the YIR proteins could be involved in antigenic variation. During *P. chabaudi* infection of immuno -competent and -deficient mice however, few substantial differences were detected by RT-qPCR. It is possible that some *cir* genes re-capitulated the altered *yir* profile in immuno-competent mice, but which escaped detection due to the limited primers available. Thus, RT-qPCR is not an ideal method for analysis of such a large gene family, by necessity focussing on transcription of a few *cir* genes.

Whole transcriptome analyses are invaluable for the investigation of multi-gene family expression patterns, in order to see the whole picture. For example, differences in the transcription of different *yir* sub-familes were observed during blood stage *P. yoelii* infection by RT-qPCR (Fonager et al., 2007), but not by the subsequent microarray study (Cunningham et al., 2009). Neither could correlations be identified by microarray between the pattern of *vir* transcription during intra-erythrocytic development and the previously identified VIR sub-families (Bozdech et al., 2008). These data indicate that members of the sub-families identified on the basis of sequence similarity were not co-expressed during *P. yoelii* infection or *P. vivax* intra-erythrocytic development. Instead, two waves of *vir* and *Pvtrag* transcription were detected during *P. vivax* intra-erythrocytic development (Bozdech et al., 2008), which may indicate that two functionally related VIR sub-families exist.

Differences in *pir* transcription have also been detected from microarray studies: between mice susceptible and resistant to the development of cerebral malaria (Lovegrove et al., 2006), and in mice immunized with MSP-8, thus inhibiting the invasion of mature erythrocytes (Shi et al., 2005). These results may indicate the

involvement of PIR proteins in RBC invasion, processes leading to the development of cerebral malaria, or could reflect changes in the whole transcriptome as a result of altered parasite behaviour. This illustrates the power of whole transcriptome analyses, since differences in *yir* transcription were detected without prior expectation that the expression of these genes would change. Unfortunately, due to the complex analysis required, many micro-array studies simply exclude sub-telomeric regions and the multi-gene families they encode [for example (Liew et al., 2010)], potentially meaning that differences in VSA expression are not detected.

The future application of RNA sequencing technologies will likely facilitate analysis of VSA genes from whole *Plasmodium* transcriptome studies, as gene-specific probes are not required. During their analysis of the *P. falciparum* transcriptome, Otto and colleagues were able to provide near-complete genome coverage of RNA transcripts at single base pair resolution, and detect alternative splicing events, novel transcripts and low abundance transcripts (Otto et al., 2010). Periodic gene expression patterns during the erythrocytic cycle (Le Roch et al., 2003, Bozdech et al., 2003a) were also detected using RNA sequencing (Otto et al., 2010). Comparison of the same samples by micro-array and RNA sequencing correlated well and furthermore, whole transcriptome sequencing was able to detect lower abundance transcripts giving an expanded transcriptome (Otto et al., 2010). It is anticipated that similar investigations will be possible for *P. chabaudi* once methods for improved sample preparation are defined.

P. chabaudi RNA samples contained high abundance contaminants, of both ribosomal RNA and host-derived mRNA. None of the investigated reticulocyte depletion strategies could provide enough material from an individual *P. chabaudi* infected mouse for cDNA library preparation, which requires approximately 10 μg RNA for the methods used here. Whilst pooling of samples from several individuals is acceptable for certain applications, this is not informative for analysis of *cir* transcription dynamics; since the parasite populations arising in each mouse may be diverse, with different *cir* gene transcription in different infections. Recently, methods requiring much smaller quantities of template for RNA sequencing have been described (Ozsolak et al., 2010), which may be able to resolve this problem.

It may be also be possible to specifically deplete the high abundance contaminants directly from *P. chabaudi* RNA samples, because they were predominantly rRNA and *globin* mRNA transcripts. To do this, biotinylated oligonucleotides could be designed to hybridize to each of the contaminating transcripts, which could then be removed from

the RNA samples using streptavidin-coated beads. This was done for 26 rRNA species and 32 of the most abundant mRNA transcripts, including those encoding histones, during analysis of the *P. falciparum* transcriptome (Otto et al., 2010). More recently, another approach has been published, where rRNA contamination was excluded at the stage of cDNA template preparation (Vignali et al., 2011). Here, Vignali and colleagues used a 'not so random' priming strategy for cDNA synthesis, where any hexamers that would bind to rRNA transcripts were selectively removed from the random primers before cDNA synthesis (Vignali et al., 2011). Either or both of these approaches could be used to exclude rRNA and globin mRNA contamination from *P. chabaudi* cDNA samples prior to RNA sequencing.

The application of whole transcriptome analysis methods, such as RNA sequencing, to the rodent model *P. chabaudi* will enable the investigation of parasite responses during growth in different host genetic backgrounds and under different treatments and is also likely to identify genes not previously known to be involved in such processes. For future studies, investigations should repeat the comparison of *P. chabaudi* infected immuno-competent and immuno-deficient mice to determine whether *cir* transcription is reduced in the absence of lymphocytes, as previously observed by RFLP analysis for the *yir*s (Cunningham et al., 2005). If a sub-set of *cir* transcripts were reproducibly detected at lower levels by RNA sequencing in the absence of lymphocytic selection (such as during *P. chabaudi* infection of RAG2$^{-/-}$ mice), this would strongly suggest that CIR proteins are under immune selection during the blood stages of infection.

Subsequently, analysis of *P. chabaudi* transcription during parasite passage in mice of the same strain would identify differences in the *cir* transcription pattern. Analysis of the whole *cir* repertoire from these experiments will allow determination of the extent to which *cir* transcription changes between *P. chabaudi* infected individuals, compared to the donor mouse, identifying which transcripts become dominant in each infection. If a restricted number of *cir*s were more likely to dominate *P. chabaudi* infection, these family members should provide the basis of investigations into CIR function.

Chapter 4: Generation of recombinant CIR proteins

4.1 Introduction

In order to investigate beyond the transcriptional analyses described in chapter 3 and explore the possible functions of CIR proteins, efforts were made to express recombinant CIR proteins. These proteins could then potentially be used for analyses of CIR protein structure, and as tools for investigation of CIR function.

Plasmodium genes are notoriously difficult to express as recombinant proteins because of their A+T rich nature, containing an average of 80.6% A+T in *P. falciparum* and 74.5% in *P. chabaudi* (Gardner et al., 2002, Janssen et al., 2001). This is unlike the typical A+T content, closer to 50%, found in common expression systems, and often leads to premature transcription termination during heterologous expression of *Plasmodium* genes [for example (Yadava and Ockenhouse, 2003)]. The *pir* genes are no exception, and the presence of at least one trans-membrane domain and other hydrophobic regions make them less likely to be soluble if expression is achieved.

Insoluble proteins form irreversible aggregates, but soluble proteins can also form aggregates, such as inclusion bodies (Carrio and Villaverde, 2002), which must be disrupted by denaturing conditions before the proteins can undergo solubilization and re-folding. After such treatment, proteins may not accurately represent the native protein structure. The structure of recombinant proteins should be equivalent to the native protein, especially for immunological investigations, where improper protein folding may result in the loss of conformational determinants that are recognised by antibodies.

Few attempts have been made to express recombinant PIR proteins. Initial endeavors by groups using *E. coli* managed with varying degrees of success to produce VIR GST-fusion proteins (Del Portillo et al., 2001, Fernandez-Becerra et al., 2005, Oliveira et al., 2006). Whilst all of these were insoluble and not necessarily representative of the endogenous protein structure, nonetheless some VIR proteins were recognized by Abs present in immune sera from *P. vivax* infected patients by western blot, indicating that non-conformational determinants of Ab specificity were present.

To enhance the likelihood of expressing soluble CIR proteins, a eukaryotic protein expression system was used for this study, *Pichia pastoris*. The yeast *P. pastoris*

contains chaperones to facilitate protein folding and machinery required for post-translational modifications. In this system it is also possible to secrete recombinant proteins directly into the culture medium, reducing lengthy protein purification strategies. Our laboratory has previously successfully expressed the 21 kDa *P. chabaudi* equivalent of *P. falciparum* $MSP1_{19}$ using this expression system (Hensmann et al., 2004).

The objectives of the work described in this chapter were to:

i) Design synthetic *cir* genes, removing the trans-membrane domain and predicted intracellular domains in concert with gene re-codonisation for optimal expression in the methylotropic yeast *P. pastoris.*

ii) Generate *P. pastoris* clones capable of expressing recombinant CIR proteins, to determine optimimal expression conditions for each recombinant CIR and to purify the recombinant CIRs via metal-chelate chromatography.

iii) Conduct initial analyses of protein structure to assess the likelihood of recombinant CIRs having equivalent protein structures to native CIRs, and the feasibility of crystallisation of these proteins.

4.2 Methods

4.2.1 Cloning and expression of CIR

The DNA sequences of three *cir* genes chosen for protein expression: PCHAS_000100, PCHAS_070130 and PCHAS_040110 were downloaded from Genedb (http://www.genedb.org/Homepage/Pchabaudi). The predicted TM and intracellular domains were removed and restriction sites were inserted to flank the remaining sequences corresponding to the predicted extracellular region. The nucleotide sequences were further modified by re-codonization for optimal expression in *Pichia pastoris* and addition of a 3'-tag encoding six-histidine residues, which would facilitate protein purification. In addition, asparagine residues predicted by the NetNGlyc 1.0 Server to undergo N-glycosylation in *P. pastoris* (http://www.cbs.dtu.dk/ services/NetNGlyc/), were altered either to amino acids found at the same position in other CIR sequences or to amino acids with similar properties, since N-glycosylation occurs only rarely in *Plasmodium* (Gowda and Davidson, 1999).

Following synthesis by GeneART™, the *cir* genes were sub-cloned into EcoRI or AvrII and NotI sites of the expression vector pPIC9K (Invitrogen). This vector contains the α-factor mating signal sequence to enable protein expression directly into the culture medium, located directly 5' of the multiple cloning site. Briefly, restriction digestion was carried out using either EcoRI or AvrII and NotI in buffer 3 (New England Biosciences), for 2 hours at 37 ˚C, according to manufacturer's instructions. The pPIC9k vector was digested in the same way, in parallel, and digested products were size separated by agarose gel electrophoresis. Bands corresponding to the synthetic *cir* gene and the pPIC9K vector backbone were excised and purified using a Qiaquick gel extraction kit (Qiagen), according to manufacturer's instructions. These products were then ligated together using the rapid ligation kit (Roche), according to manufacturer's instructions and transformed into TOP10F *E. coli* cells (Invitrogen), as described previously (chapter 3.2.5). Several transformants were selected for a repeat double-restriction digestion to confirm which cells contained the full *cir* gene and vector. A clone containing the expected *cir* gene was then inoculated into 500 ml cultures of LB broth (1% tryptone, 0.5% yeast extract, 1% NaCl, pH 7.0) containing 50 μg/ml ampicillin, and cultured overnight, shaking, at 37 ˚C. Plasmids containing the synthetic *cir* genes were then purified from the culture by maxiprep (Qiagen), according to manufacturer's instructions. Plasmids were eluted in distilled water and their

concentrations were estimated using a Nanodrop spectrophotometer (ThermoFisher). The plasmids were sequenced using M13 primer sites flanking the *cir* gene insertion, for confirmation that the expected gene was present and in frame with the α-factor signal sequence.

Electroporation of GS115 *P. pastoris* cells (Invitrogen) was carried out according to methods adapted from the pPIC9k manual (Invitrogen). Between 1 and 5 µg of linearized vector was added to 60 µl *P. pastoris* cell suspension (containing approximately 1.2 x 10^9 cells, re-suspended in buffer comprising 270 mM sucrose, 10 mM Tris-HCl, 1 mM $MgCl_2$, pH 7.5), in an 0.2 cm electroporation cuvette (BioRad). This was placed in the gene pulser (BioRad) and electroporated using 1.5kV, 400Ω, 25µF. Recombinants were selected by growth firstly on minimal medium- (containing 1.34% yeast nitrogen base, $4x10^{-5}$% biotin and 2% dextrose), which selected for complementation of histidine auxotrophy in the GS115 *P. pastoris* cells, by the histidine biosynthesis open reading frame (ORF) present in the vector pPIC9k. Secondly, recombinants were selected by growth on yeast extract peptone medium- (containing 1% yeast extract, 2% peptone and 2% dextrose) agar plates containing increasing concentrations of Geneticin ® (Invitrogen) selection. This allowed selection of recombinants with high Geneticin ® resistance, which correlates with an increased copy number of integrated constructs, thereby leading to higher levels of recombinant protein expression (Clare et al., 1991a).

Clones growing at the highest concentration of Geneticin ® (1 mg/ml) were isolated, and used to create glycerol stocks, made by addition of 0.15 ml glycerol to 0.85 ml of *P. pastoris* culture, grown in yeast extract peptone medium. These were 'snap frozen' on dry ice for storage at -80 ˚C. In addition, the clones were transferred to BMGY medium (liquid broth containing 1% yeast extract, 2% peptone, 100 mM potassium phosphate pH 6.0, 1.34 % yeast nitrogen base, $4x10^{-5}$ % biotin and 0.5 % glycerol); and cultured at 28 ˚C, in a shaking incubator set at 225 rpm, until the OD_{600} reached 2-6, indicating log phase growth.

P. pastoris is able to use methanol as a sole carbon source, the first step in this pathway is the oxidation of methanol to formaldehyde using the enzyme alcohol oxidase. The enzyme responsible is produced by the *AOX1* gene, whose promoter is encoded within the vector pPIC9k (Invitrogen), where it is fused to the α-factor secretion signal (Ellis et al., 1985, Koutz et al., 1989, Tschopp et al., 1987). Growth using methanol as the sole carbon source requires induction of the *AOX1* gene (Cregg et al., 1989), and therefore

also induces expression of the recombinant *cir* gene inserted downstream of the α-factor signal sequence.

Recombinant protein expression was induced by re-suspension of the cells in BMMY medium, where glycerol was replaced with methanol as the sole carbon source, according to methods adapted from the pPIC9k manual (Invitrogen). Cells were otherwise cultured in the same conditions as described above, adding a further 0.5% v/v methanol to maintain protein expression after every 24h of culture. For standard trials of protein expression, cells were re-suspended in BMMY to an OD_{600} of 1, in order to maintain log phase growth. Protein expression was optimized by addition of 5 μg/ml chymostatin (Sigma) every 24 h or re-suspension of *P. pastoris* in minimal methanol medium (Liquid broth containing 1.34% yeast nitrogen base, $4x10^{-5}$% biotin and 0.5% methanol) instead of BMMY. Both strategies inactivate host proteases, thus preventing the degradation of full-length recombinant proteins during protein expression (Clare et al., 1991b, Brierley et al., 1994). To favour slow methanol metabolism in clones where the *AOX1* gene may have been replaced (resulting in poor utilisation of methanol as the sole carbon source) cells were re-suspended in BMMY at 20% of the original culture volume.

4.2.2 Purification of recombinant CIR

Recombinant CIR proteins containing hexa-histidine tags were purified by immobilised metal-affinity chromatography on Ni^{2+}-nitrilotriacetic acid (Ni^{2+}-NTA) agarose (Invitrogen), under native conditions, according to methods adapted from the Ni^{2+}-NTA manual (Invitrogen). Briefly, 45 volumes of culture supernatant were allowed to bind one volume of Ni^{2+}-NTA agarose in the presence of complete protease inhibitors (Roche) and five volumes of binding buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole, pH 8.0) for one hour, rotating, at 4 ˚C. Non-specifically bound proteins were removed using 25 volumes of wash buffer (50 mM NaH_2PO_4, 300 mM NaCl, 20 mM imidazole, pH 8.0), and bound protein eluted with three volumes of elution buffer (50mM NaH_2PO_4, 300 mM NaCl, 200 mM imidazole, pH 8.0). Centrifugation steps were carried out at 4 ˚C, using 3,000 x *g* for volumes greater than 1.5 ml and 13,000 x *g* for volumes smaller than 1.5 ml.

During the initial screens of protein expression, proteins were eluted from the Ni^{2+}-NTA agarose using one volume of three elution buffers, each containing a higher imidazole concentration: 250 mM; 500 mM and 1 M imidazole, to determine whether all the protein had been eluted.

Protein concentration was analysed using the bicin-choninic acid assay (BCA, Pierce), with bovine serum albumin (BSA) diluted in the appropriate buffer to form a standard curve. The absorbance of samples at 562 nm was measured using a Nanodrop spectrophotometer (Thermo Scientific). For subsequent studies, recombinant protein was dialyzed into PBS using either Dialyzer cassettes (Pierce) or desalted into PBS using PD10 columns (Amersham Biosciences).

4.2.3 CIR detection via SDS-PAGE and immuno-blotting

Time-course experiments were carried out, consisting of one ml samples collected every 24 hours during induction of protein expression, which were analysed by 12% sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE). Concentration of all protein samples was necessary, achieved by centrifugation at 13,000 x *g*, 4 ˚C, using Vivaspin concentrating columns with a MWCO pore size of 10 kDa (Vivascience), until the volume of each sample was reduced ten-fold.

Proteins were resolved on NuPAGE 12% bis-Tris gels (Invitrogen) in NuPAGE 1x morpholine-ethane-sulfonic acid sodium dodecyl sulphate buffer (Invitrogen), according to the manufacturer's directions. Reducing conditions were used throughout, where samples contained 100 mM DTT and the running buffer also contained 0.5 ml NuPAGE antioxidant. Proteins were run alongside 1x SeeBlue Plus2-prestained standard (Invitrogen). Gels were either directly stained with coomassie blue or used for immuno-blotting. Coomassie blue staining was carried out overnight (0.25% w/v coomassie brilliant blue R250, 45% ethanol, 10% acetic acid) and gels were subsequently destained for several hours (45% ethanol, 10% acetic acid). Once fully destained, gels were washed in distilled water, and dried using the GelAir system (BioRad).

For immuno-blotting, the proteins were electrophoretically transferred to Hybond C membrane (Amersham Biosciences) in buffer containing 10% methanol, 0.025 M Tris base (Sigma) and 0.2 M Glycine (BDH Biosciences) for 3 h at 30 V. Membranes were

blocked at 4 ˚C overnight in Odyssey blocking buffer (Licor Biosciences), and washed for 10 minutes three times with PBS containing 0.05% Tween 20. Specific proteins were detected using the anti-six his tag monoclonal antibody (Novagen) at 0.5 µg/ml concentration, membranes were washed as before. Bound antibodies were detected using secondary Alexa 680-conjugated goat anti-mouse IgG (Licor Biosciences) used at 0.13 µg/ml concentration. Membranes were washed as before and stored in PBS in the dark. The bound secondary antibody was detected by scanning the probed membranes in the Odyssey scanner (Licor Biosciences), exciting the Alexa flurochrome at 680nm, and detected at 700nm.

4.2.4 Biochemical measurements of recombinant CIR

Dynamic light scattering was carried out with 20 µl protein samples, using a Viscotek spectrophotometer in conjunction with Omnisize 3.0 software.

Circular dichroism and aromatic absorbance spectra were carried out as follows:

- Uncorrected fluorescence emission spectra were recorded using a Jasco FP-6300 spectrofluorimeter equipped with an ETC-273T temperature controller with excitation at 280 nm (bandwidth 1.5 nm) and emission scanned from 290 to 400 nm (bandwidth 5 nm). All measurements were made at 20 ^{o}C in Tris HCl pH 7.2, to which 500 mM L-proline (Sigma) was added for one sample of PCHAS_000100 to enable concentration of the protein [the use of proline for this purpose has been reviewed by (Alibolandi and Mirzahoseini, 2011), and used for example by (Schobert and Tschesche, 1978)].
- Far-UV CD spectra were recorded on a Jasco J-715 spectropolarimeter equipped with a PTC-348WI temperature controller. All measurements were made in Tris HCl pH 7.2 at 20 ^{o}C using 1 mm quartz cuvettes and appropriate baselines were subtracted. Spectra are presented as the CD absorption coefficient calculated on a mean residue weight basis (De_{MRW}). Secondary structure content was estimated using previously described methods (Sreerama and Woody, 2000).

4.3 Results

4.3.1 Selection of *cir*s for recombinant protein expression

Three *cir* genes were chosen for expression in *P. pastoris* according to their transcription level during blood stage *P. chabaudi* infection. Transcription of PCHAS_000100 and PCHAS_040110 was identified by RT-qPCR, using primers designed to amplify these genes (described in chapter 3.3.1), and all three *cir*s were also identified from the trial Illumina RNA sequencing experiments (described in chapter 3.3.2). PCHAS_070130 was highly transcribed in both samples used for RNA sequencing.

The locations of the chosen *cir* genes within the whole family are shown within a network in Figure 27a. The design of synthetic *cir* genes is summarized in Figure 27b, as a flow chart, i), and with the DNA sequences for each synthetic *cir* also shown, ii). The encoded amino acid sequences of all synthetic *cir* genes are shown in Figure 28a. These three recombinant CIRs had different predicted protein characteristics, for example, Figure 28b shows the hydrophobicity profile of each protein (Gasteiger et al., 2005). This indicated that PCHAS_070130 contained more hydrophobic residues than the other CIRs, which may affect solubility of the expressed protein. In addition, the predicted protein characteristics are shown in Table 19, calculated using Expasy ProtParam (Gasteiger et al., 2005). The instability index is based upon the presence of particular dipeptides found in proteins known to be degraded in solution, where a score greater than 40 indicates the protein is likely to be degraded (Guruprasad et al., 1990). Only PCHAS_070130 was predicted to be stable so this protein was expected to be less prone to degradation than the other CIRs.

Table 19: Statistics predicted for each recombinant CIR protein.

CIR protein	Amino acid no.	Molecular weight (Da)	Isoelectric point	1] Extinction coefficient [1.]	[2.]	Instability index	2] GRAVY
Pc_000100	239	28758.2	5.46	41510	41260	47.99 'unstable'	-0.964
Pc_070130	257	29361.2	5.91	28350	27850	38.67 'stable'	-0.400
Pc_040110	284	32871.1	6.07	42580	41830	44.86 'unstable'	-0.6888

The species identifier PCHAS_ was abbreviated to Pc_. All predictions were made using ProtParam (Gasteiger et al., 2005). 1] Extinction coefficients (M^{-1} cm^{-1} at 280 nm in H_2O) were calculated on the basis that either [1.] all cysteine residues appear as ½ (ie all participate in disulphide bonding), or [2.] no cysteine residues appear as ½. 2] The 'grand average of hydropathy' was abbreviated to GRAVY.

4.3.2 Expression of recombinant CIR

Figure 29 shows the cloning strategy employed for generation of recombinant *Pichia*, and the resulting maps of cloned synthetic *cir* sequences within the pPIC9k vector. The flow-diagram depicted in Figure 29c was followed to generate recombinant *P. pastoris* clones. The clones that grew on high Geneticin ® concentration plates were assumed to have integrated more copies of the vector and so would be likely to express higher yields of protein. Selected clones were screened for protein expression (data not shown) and the highest expressers were used to determine the optimal length of protein expression.

Figure 30 shows a time-course of protein expression for the three synthetic CIRs. Coomassie stained SDS-PAGE gels show the presence of bands at the expected size of approximately 29 kDa for PCHAS_000100 and PCHAS_070130. Bands above 40 kDa in size were also visible for both proteins, appearing at higher concentration for PCHAS_070130. The composition of these high molecular weight bands was unknown. PCHAS_000100 and PCHAS_070130 were the expected size, at 29 kDa, although PCHAS_070130 appeared as a doublet along with a slightly higher molecular weight band. PCHAS_070130 clearly aggregated forming high molecular weight bands visible by coomassie blue staining of SDS-PAGE gels, which were no longer recognized by anti-his antibody in western blots, Figure 30b. The presence of a doublet at the expected size and slightly larger may be the result of post-translational modifications that caused the protein to migrate differently in SDS-PAGE. Since N-glycosylation sites were removed from the synthetic CIR sequences, a likely post-translational modification of PCHAS_070130 is phosphorylation. Because of time constraints, this was not investigated further.

By contrast, PCHAS_040110 expression resulted in a smaller than expected protein, Figure 30. The PCHAS_040110 observed was half the expected size, at 16 kDa, and the fragment was not recognized by the anti-his antibody in Western blot. Either the secreted protein was not CIR PCHAS_040110, or it was PCHAS_040110 but degraded or not fully transcribed, such that the C-terminal his tag was missing. To determine which of these was the case, protein expression was induced in a range of conditions to inactivate yeast proteases, shown in Figure 31. In parallel, samples were taken from an un-induced culture, to determine whether the 16 kDa protein was also expressed in the absence of methanol induction, and therefore likely to be a secreted host protein. This

was not the case. The full size PCHAS_040110 protein could be obtained by expression under conditions which inactivated neutral pH host proteases (Brierley et al., 1994), indicating that protease degradation led to the smaller than expected product. Specifically these conditions used minimal, un-buffered culture media and BMMY culture media containing the protease inhibitor chymostatin (Sigma).

4.3.3 Purification and yield of recombinant CIR

All three recombinant CIRs could be purified using metal chelate chromatography, producing clean proteins of the expected sizes (PCHAS_000100 28.7 kDa, PCHAS_070130 29.3 kDa and PCHAS_040110 32.8 kDa), Figure 32. Elution of the proteins from the Ni^{2+}NTA agarose used three buffers, which contained increasing strengths of imidazole, to elute the majority of recombinant protein from the beads. PCHAS_000100 appeared to be expressed at higher levels than the other CIRs, and the majority of this protein was eluted using the lowest imidazole concentration, 250 mM. By contrast, PCHAS_070130 and PCHAS_040110 had lower protein yields to start with, and these proteins were most successfully eluted with the highest strength imidazole, 1M. All purified CIR proteins also contained high molecular weight products, which were also recognized by the anti His-tag Abs via Western blot. These could have been aggregated CIR proteins, as their molecular weights were approximately double the size of the recombinant CIRs.

The protein yields were estimated from large-scale expression and purification of the recombinant proteins. This was carried out to predict the culture quantity required for protein production and use in subsequent assays. Optimal protein yields from 100 ml culture are shown in Figure 33a. Both PCHAS_070130 and PCHAS_040110 were expressed at much lower levels than PCHAS_000100, producing approximately 0.3, 0.5 and 5 mg protein from 100 ml culture, respectively. Therefore, protein yields under different culture conditions were compared. For PCHAS_040110, growth in minimal medium or BMMY containing 5 μg/ml chymostatin was compared (Figure 33b), where the latter method produced a higher yield of recombinant protein measured by BCA assay. For PCHAS_070130, there was no suggestion of proteolytic cleavage since the full-length product was expressed in BMMY media without the addition of protease inhibitors. Low protein yields may have been a result of impaired methanol metabolism by the *P. pastoris* clone, for example if the *AOX1* methanol utilization gene was

deleted. This was not the reason for poor protein expression, since no difference in protein yield was observed after growth in conditions favourable to methanol-slow (Mut S) or wild type (Mut +) *P. pastoris* (Figure 33c).

4.3.4 Behaviour of recombinant CIR proteins in solution

As one of the aims of expressing recombinant CIR proteins was to investigate their structural characteristics, measurements were taken to determine whether crystallization of the proteins would be feasible, to allow X-ray crystallographic determination of their structures.

Spectroscopy techniques can be used to determine certain features of proteins in solution, such as dynamic light scattering (DLS). This technique uses spectrophotometry to analyse protein particles as they move through solution [reviewed by (Dahneke, 1983, Podzimek, 2011)]. When the laser is deflected by a protein particle, the Brownian motion of particles causes a fluctuation in the scattering intensity over time. From these measurements, the average particle radius may then be determined, and the molecular weight calculated. Thus it can be determined whether the proteins exist as monomers, dimers or aggregates in solution. DLS was carried out for all three proteins to determine whether it would be possible to crystallize them for structural studies.

Figure 34a shows an example of DLS analysis using PCHAS_000100. First, OmniSIZE 3.0 (Viscotek) software calculated a baseline of radius counts using 100 measurements. Ten measurements were then taken, and used to calculate the results and error (panel i); the combined correlation function (panel ii) was the average from these readings. The residuals of the combined correlation function (plotted in panel iii against time) showed an even distribution both above and below zero, indicating that the solution was equally dispersed. The mass distribution (panel iv) showed that the major particle in solution had a radius of 2.5nm, which equated to a molecular weight of 29.58 kDa. This was correct for the recombinant PCHAS_000100 CIR protein, indicating that in PBS this protein existed as a monomer in solution.

Table 20 summarizes the DLS measurements taken from several different preparations of each recombinant protein. The number of peaks detected indicates how many protein species were present in solution. When combined with the SD and residuals of the

combined correlation function, this gives a measure of how well dispersed the protein was in solution. Finally, the molecular weight of each protein species is given, allowing the determination of whether each protein species was present as a monomer, aggregate, or as degraded material.

Neither PCHAS_040110 nor PCHAS_070130 samples contained proteins of the expected molecular weight. Instead all samples measured for these proteins included small particles; likely degradation products, and large molecular weight aggregates, which probably comprised both full length and degraded CIR proteins. Protein solutions containing a mixture of species are unsuitable for determination of protein structure, hence PCHAS_000100 was the only protein preparation for which structural analyses could be undertaken.

Table 20: Dynamic light scattering measurements of the three recombinant CIRs.

Concentration (mg/ml)	Peak	% Area	SD	% RSD	MW (kDa)
PCHAS_000100:					
0.79 *	1	100	0.23	9.2	29.58
2.16 *	1	100	1.06	30.4	62.53
7.25 +	1	100	18.24	48.7	1.70×10^4
PCHAS_070130:					
0.44	1	93.1	0.15	6.9	20.47
	2	6.9	14.37	55.7	7175.03
0.56	1	91.2	0.00	0.0	0.02
	2	8.4	0.08	6.1	7.02
	3	0.4	0.76	9.9	408.67
PCHAS_040110:					
0.12	1	75.0	0.02	3.4	1.22
	2	21.2	0.56	10.2	181.60
	3	3.8	16.28	43.6	1.49×10^4
0.94	1	95.4	1.07	35.3	47.27
	2	4.6	13.23	59.8	4987.14

DLS measurements of the three recombinant CIR proteins, taken using different protein preparations: prepared either in 1 M imidazole buffer, * in PBS, or + in PBS containing 0.5 M Proline. For each protein preparation, between 1 and 3 peaks could be measured, where peak refers to the number of species observed in solution. Area is the proportion of each species in the total sample, SD is standard deviation, % RSD is the percentage of residuals and MW is the molecular weight of each species.

Unlike the other recombinant CIRs, PCHAS_000100 was completely monomeric in solution, at low concentration (less than 1 mg/ml, Figure 34), suggesting that it likely exists as a monomer *in vivo*. Unfortunately, as the protein preparations were concentrated, the size of the protein particles increased, indicating aggregation. The formation of aggregates could be due to multimerization or stochastic particle combination, the latter of which would be incompatible with further structural characterization by X-ray crystallography.

4.3.5 Conformation of recombinant PCHAS_000100 supports bio-informatic prediction

No monoclonal antibodies have been generated against native CIR proteins, hence few methods were available to determine whether the recombinant proteins were correctly folded. One indicator was to assess the secondary protein structure by circular dichroism spectrometry (CD), and compare this to predicted secondary protein fold analyses.

CD analysis of a low concentration sample of PCHAS_000100, Figure 35a, revealed that the secondary structure contained 41.8% alpha helices and 11.8% beta sheets, with linking elements comprised of 16.9% tight turns and 28.3% random coils. PCHAS_000100 samples of higher concentration required the addition of 0.5 M proline to maintain protein solubility, which was unfortunately incompatible with CD analysis.

To compare the PCHAS_000100 samples of both low and high concentration, measurements of the aromatic spectrum were made instead of CD, Figure 35b. These results were highly similar, both fluorescence spectra having maximal emission at approximately 324 nm, signifying that both the samples were folded and that the tryptophan residue was deeply buried in the hydrophobic core [solvent exposed tryptophan has an emission maximum at 356 nm, explained by: (Shirley, 1995)]

Together, these results indicated that PCHAS_000100 was folded similarly to the native protein, as the percentage of random coil measurements was low and the tryptophan residue was not solvent-exposed. In addition, the fact that both samples retained highly similar fluorescence spectra suggests that the secondary structure of this protein was not impaired even when the protein existed in high molecular weight particles, Figure 34.

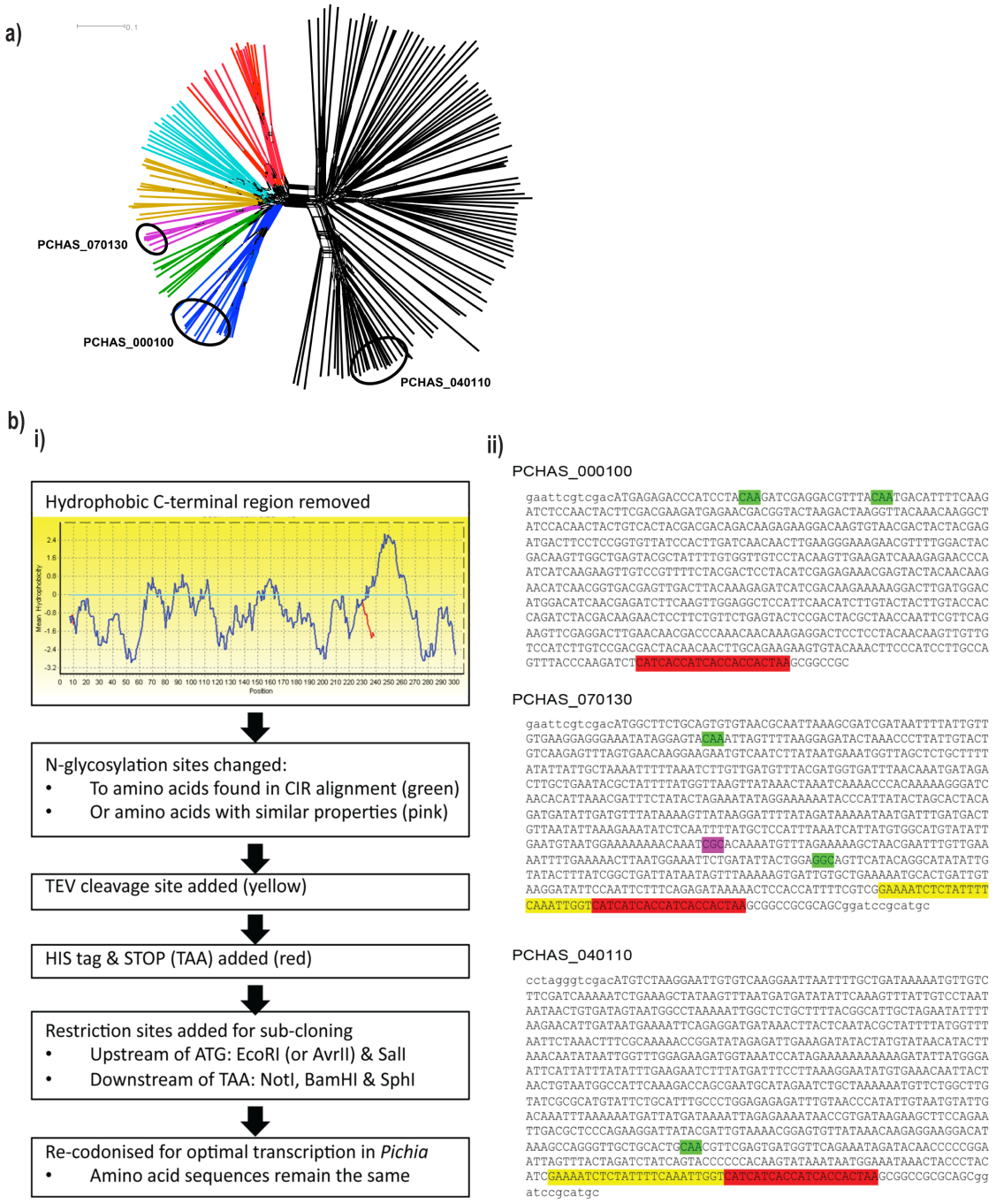


Figure 27: Choice of *cir*s for synthetic gene design.

a) Three highly transcribed *cirs* were chosen for synthetic gene design, for recombinant protein expression. Their locations within the CIR family PhyML tree are highlighted with circles.

b) Synthetic gene design. **i)** Flow chart describing the criteria applied to each *cir* gene. A mean hydrophobicity profile was created for each translated *cir* gene (Kyte and Doolittle, 1982), with a scan window of 13, in the program Bioedit 7.0.9.0 (Hall, 1999). In blue is shown the hydrophobicity profile of the orginal CIR, and in red, the profile after design of the synthetic *cir* gene PCHAS_000100. **ii)** The DNA sequences of each synthetic *cir*, before re-codonisation. These sequences were sent to GeneArt™ for re-codonisation and synthesis.

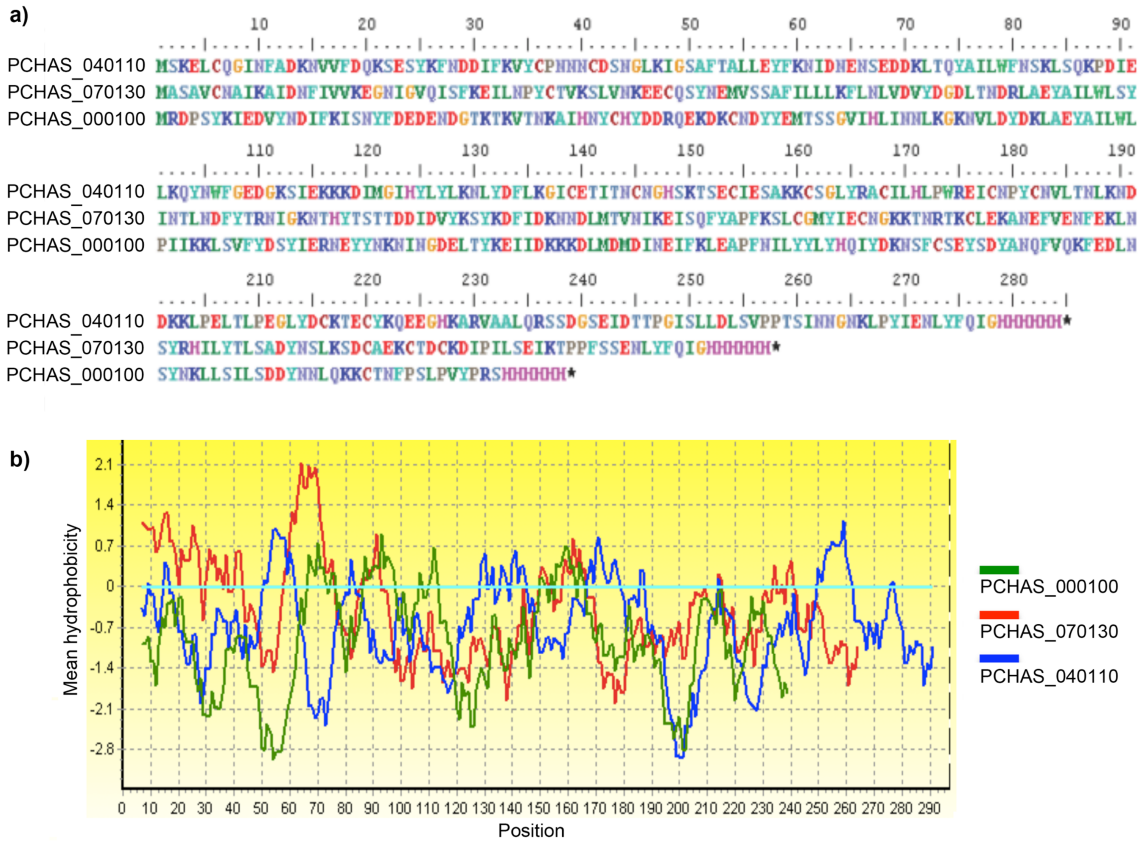


Figure 28: Comparison of the three synthetic CIR.

a) The amino acid sequences of the three synthetic CIRs obtained from GeneArt. Sequences are shown unaligned to enable comparison of the protein length.

b) The hydrophobic regions of each translated *cir* gene were determined using a Kyte and Doolittle mean hydrophobicity profile (Kyte and Doolittle, 1982), with a scan window of 13, in the program Bioedit 7.0.9.0 (Hall, 1999).

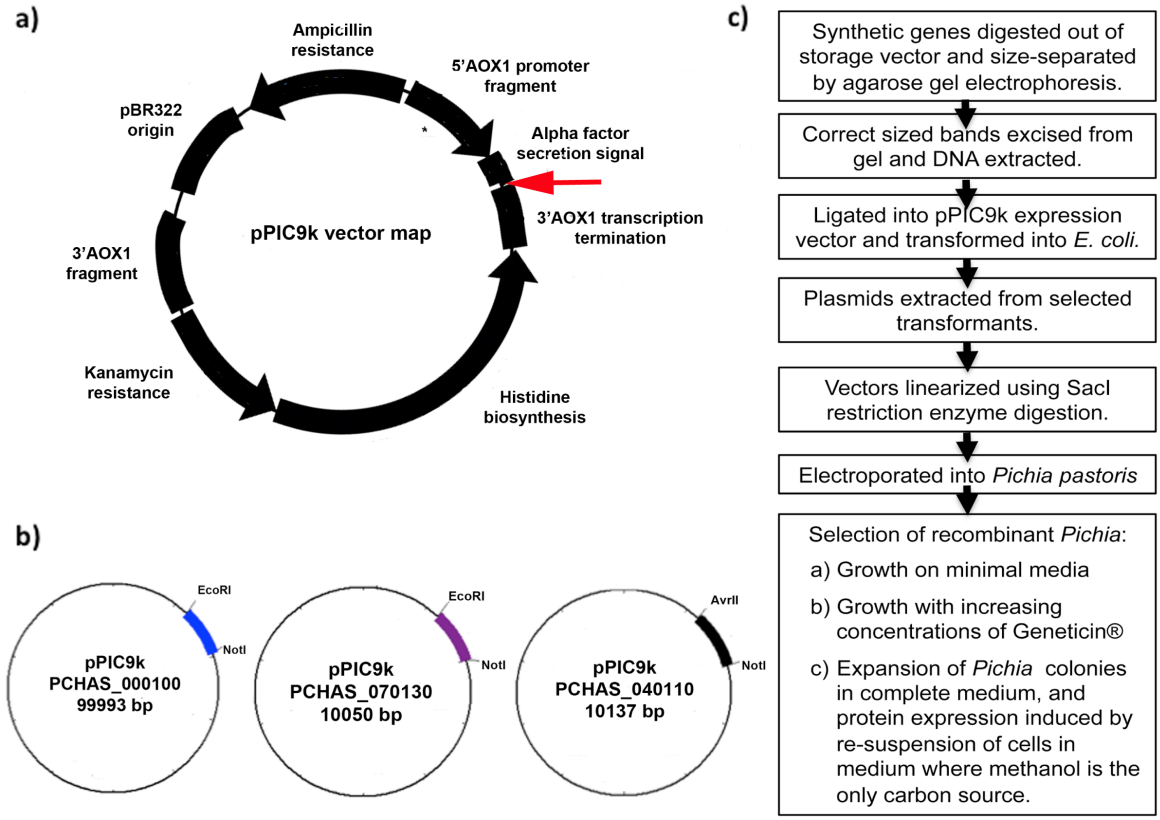


Figure 29: Synthetic *cir* cloning strategy.

a) Vector map of the features of pPIC9k (Invitrogen). * represents the SacI linearization site at position 209. The multiple cloning site is indicated by a red arrow.

b) Vector maps of the synthetic *cir* after sub-cloning into the pPIC9k vector via EcoRI or AVRII and NotI restriction sites.

c) Flow chart of the cloning strategy and selection of recombinants in *P. pastoris*.

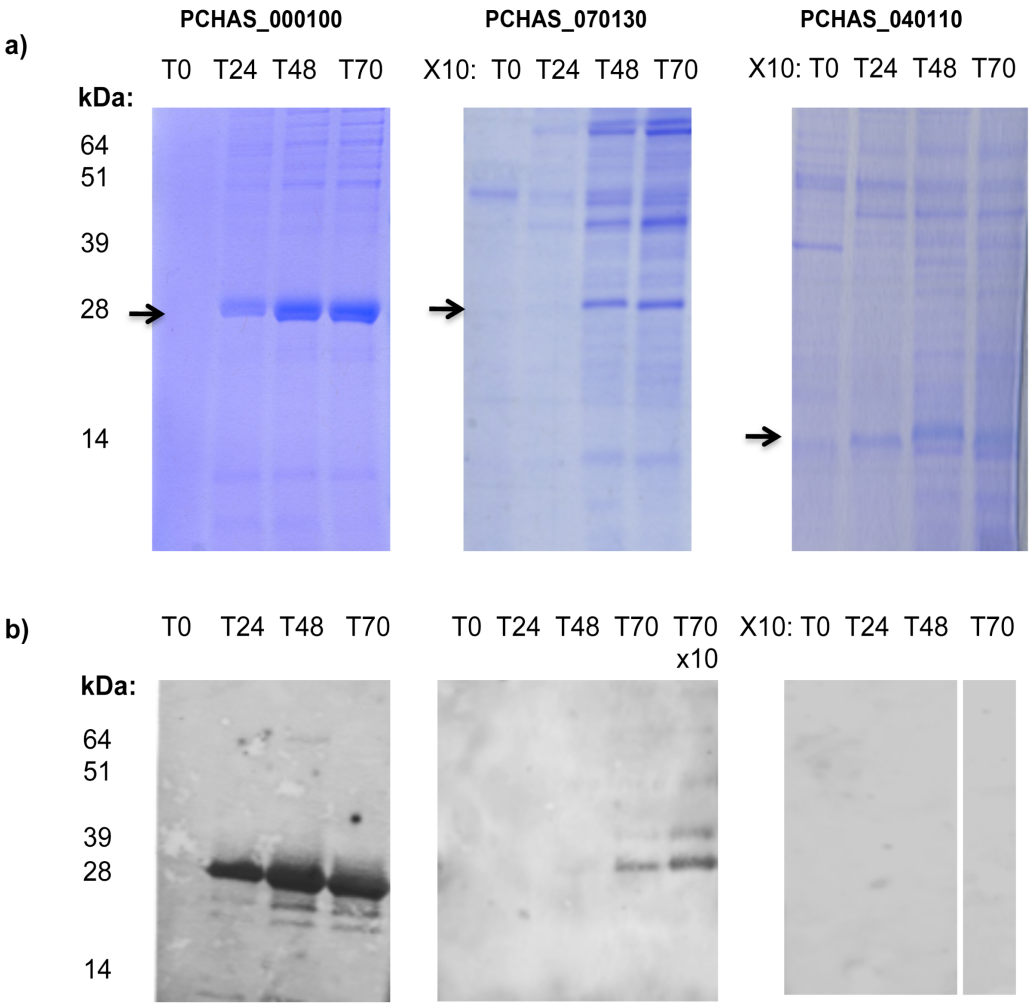


Figure 30: Recombinant CIR expression in *P. pastoris.*

Samples taken at induction (0) and at 24, 48 and 70 hours during a time-course of protein expression in *P. pastoris* for each recombinant CIR. Gels were loaded with 20 µl *P. pastoris* culture supernatant. T refers to the time in hours since induction of protein expression, x10 indicates that the sample was concentrated ten-fold prior to loading on the gel.

a) Coomassie stained SDS-PAGE gels. Arrows indicate proteins of the expected size for each recombinant CIR.

b) Immuno blot detection of the His-tags of recombinant CIRs, present in *P. pastoris* culture supernatants. Controls for the anti-His Ab are attached in appendices.

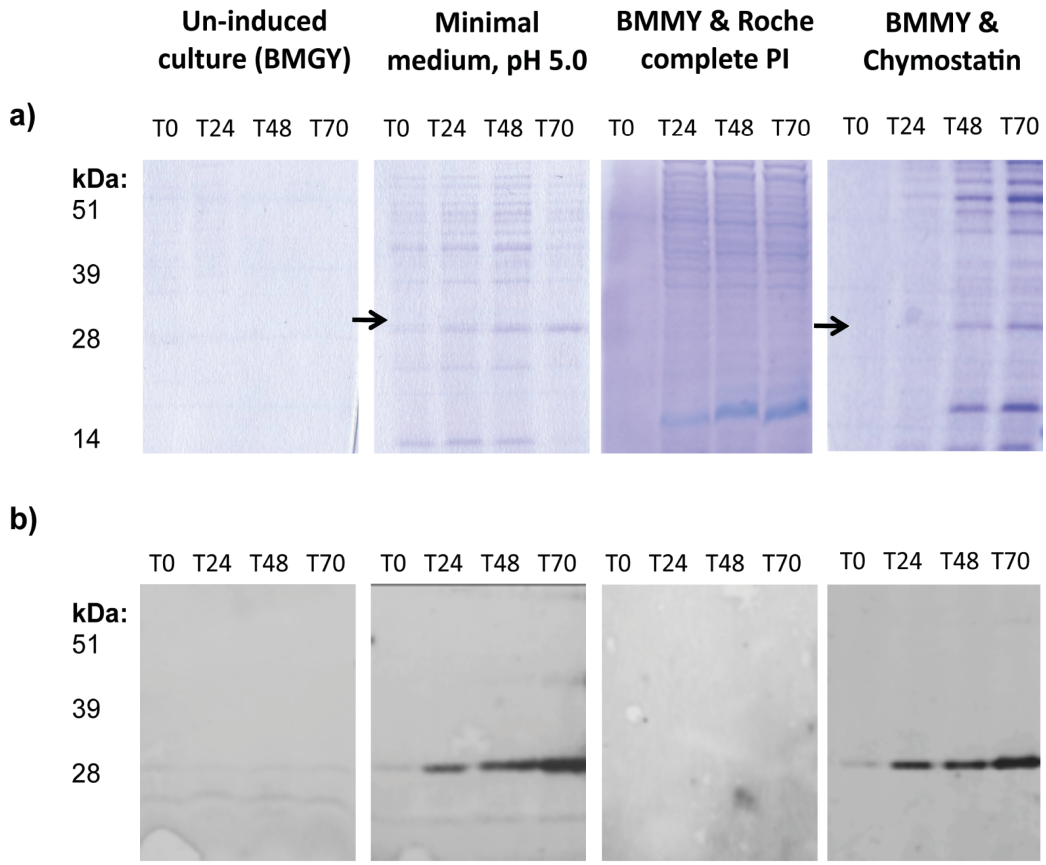


Figure 31: Optimisation of PCHAS_040110 expression.

Samples taken at induction (0) and at 24, 48 and 70 hours during a time-course of protein expression in *P. pastoris* for PCHAS_040110. A range of growth conditions was used: un-induced culture (complete 'BMGY' medium, lacking methanol), minimal medium, complete 'BMMY' medium, containing methanol plus either complete protease inhibitor tablets (Roche) added to 1% concentration every 24 h, or chymostatin (Sigma) added to 5 µg/ml concentration every 24 h.

Gels were loaded with 20 µl *P. pastoris* culture supernatant, all samples concentrated ten-fold prior to loading on the gel. T refers to the time in hours since induction of protein expression, with the exception of 'BMGY' culture which was un-induced but cultured for the same length of time.

a) Coomassie stained SDS-PAGE gels. Arrows indicate proteins of the expected size for PCHAS_040110.

b) Immuno blot detection of the His tag of PCHAS_040110 present in *P. pastoris* culture supernatants. Controls for the anti-His Ab are attached in appendices.

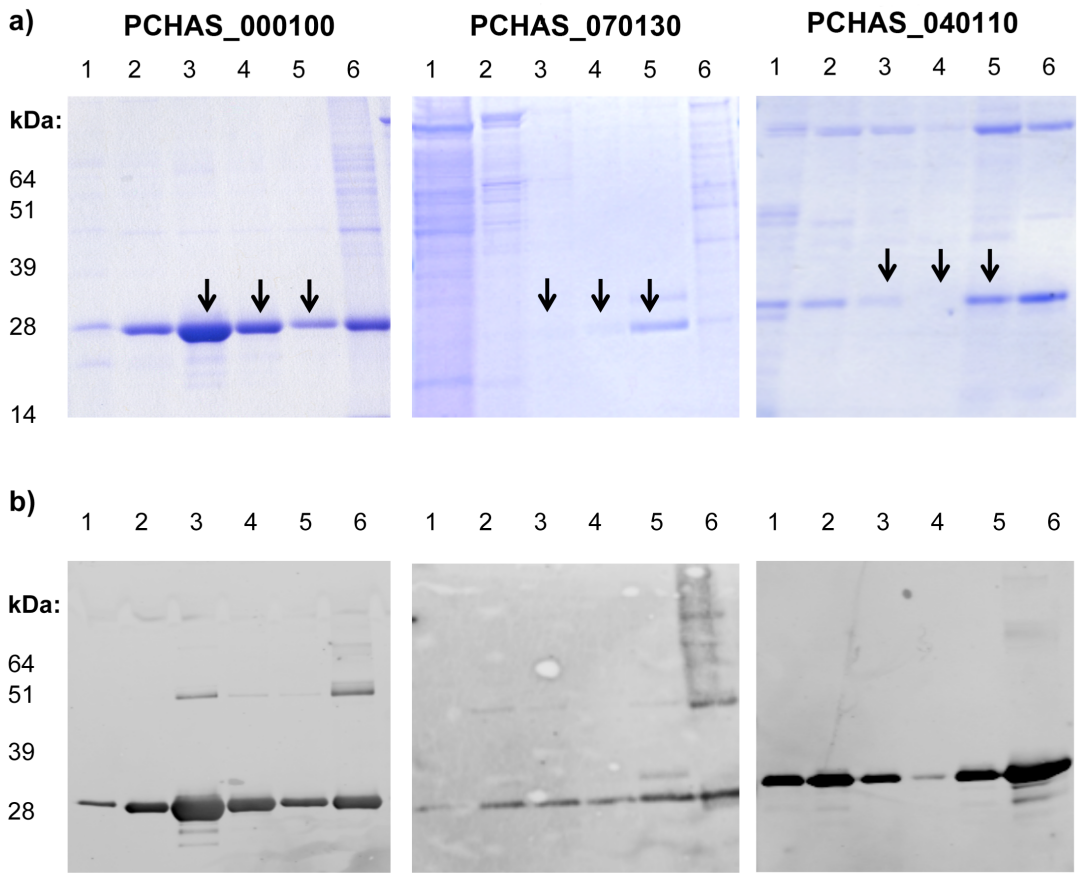


Figure 32: Purification of CIR proteins using metal chelate chromatography.

One ml samples were taken during purification of recombinant CIR proteins and concentrated ten-fold prior to SDS-PAGE, where all gels were loaded with: 1) Culture supernatant, 2) Wash, 3) Eluate using 250 mM imidazole, 4) Eluate using 500 mM imidazole, 5) Eluate using 1 M imidazole, 6) $NiNTA^{2+}$ agarose.

a) Coomassie stained SDS-PAGE gels. Arrows indicate eluted proteins of the expected size for each CIR.

b) Immuno blot detection of His tags of recombinant CIRs present after protein purification. Controls for the anti-His Ab are attached in appendices.

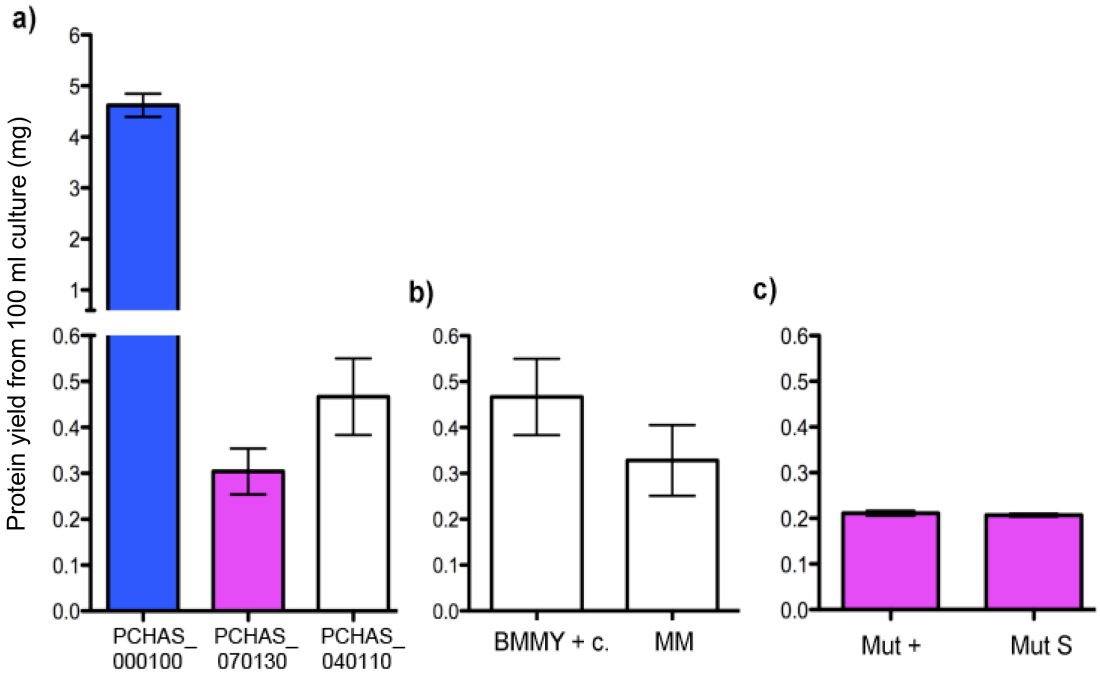


Figure 33: Comparison of protein yield for the three recombinant CIRs.

a) Estimation of protein yield from 100ml culture for each of the 3 recombinant CIRs.

b) Comparison of protein yield from PCHAS_040110 grown in either minimal medium (MM) or complete medium containing 5 µg/ml chymostatin (BMMY+c).

c) Comparison of protein yield from PCHAS_070130 grown under Mut S and Mut + conditions.

The mean protein from three replicate cultures is plotted, with error bars representing the standard error of the mean.

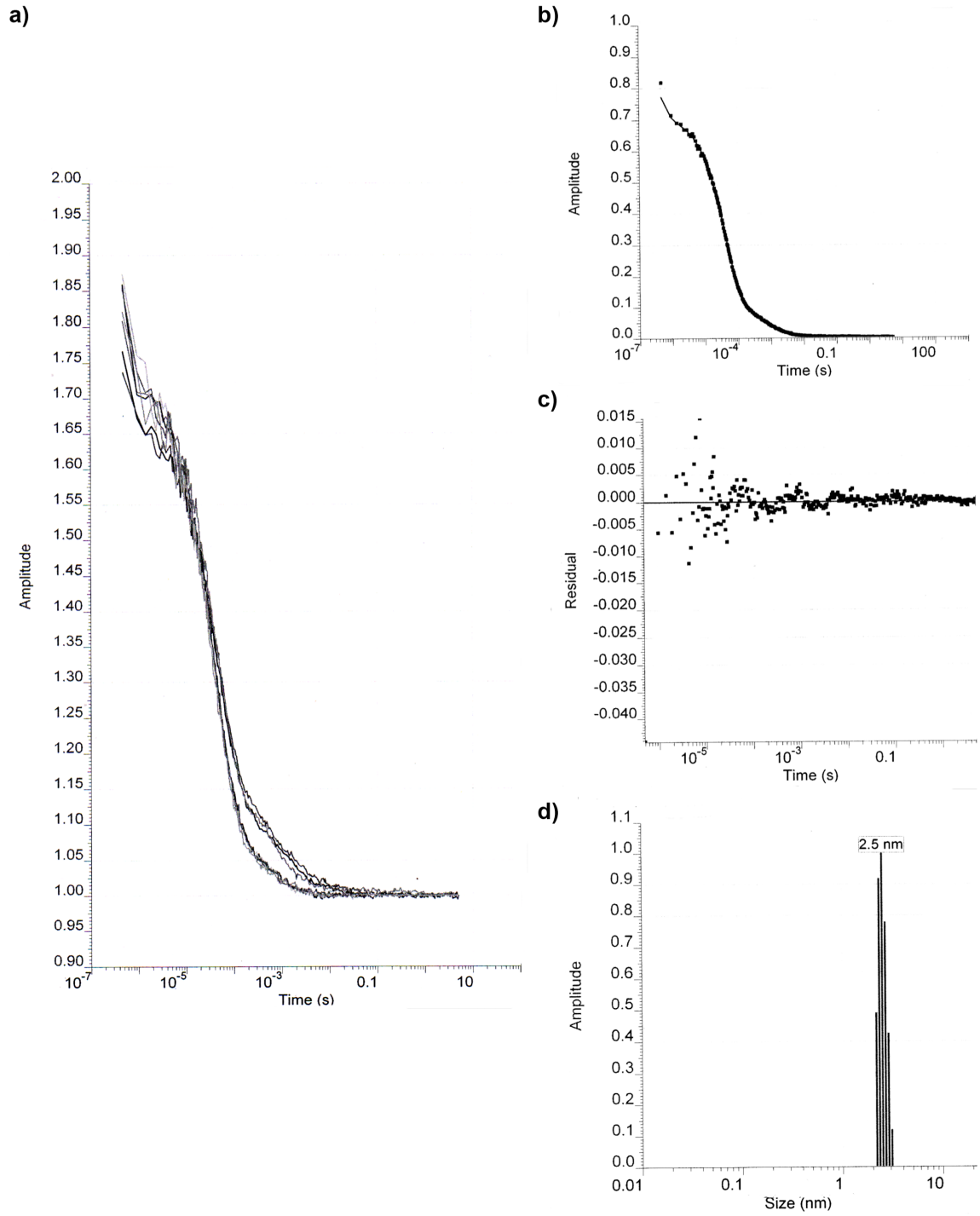


Figure 34: Measurement of PCHAS_000100 protein behaviour in solution by dynamic light scattering spectrometry.

An overview of the measurements made during DLS, using PCHAS_000100 as an example. Clockwise description of panels: **a)** The amplitude and error of ten measurements of the sample, plotted against time, **b)** The combined correlation function shows the average of these ten readings, **c)** Residuals, plotted against time, **d)** The mass distribution of the major species in solution.

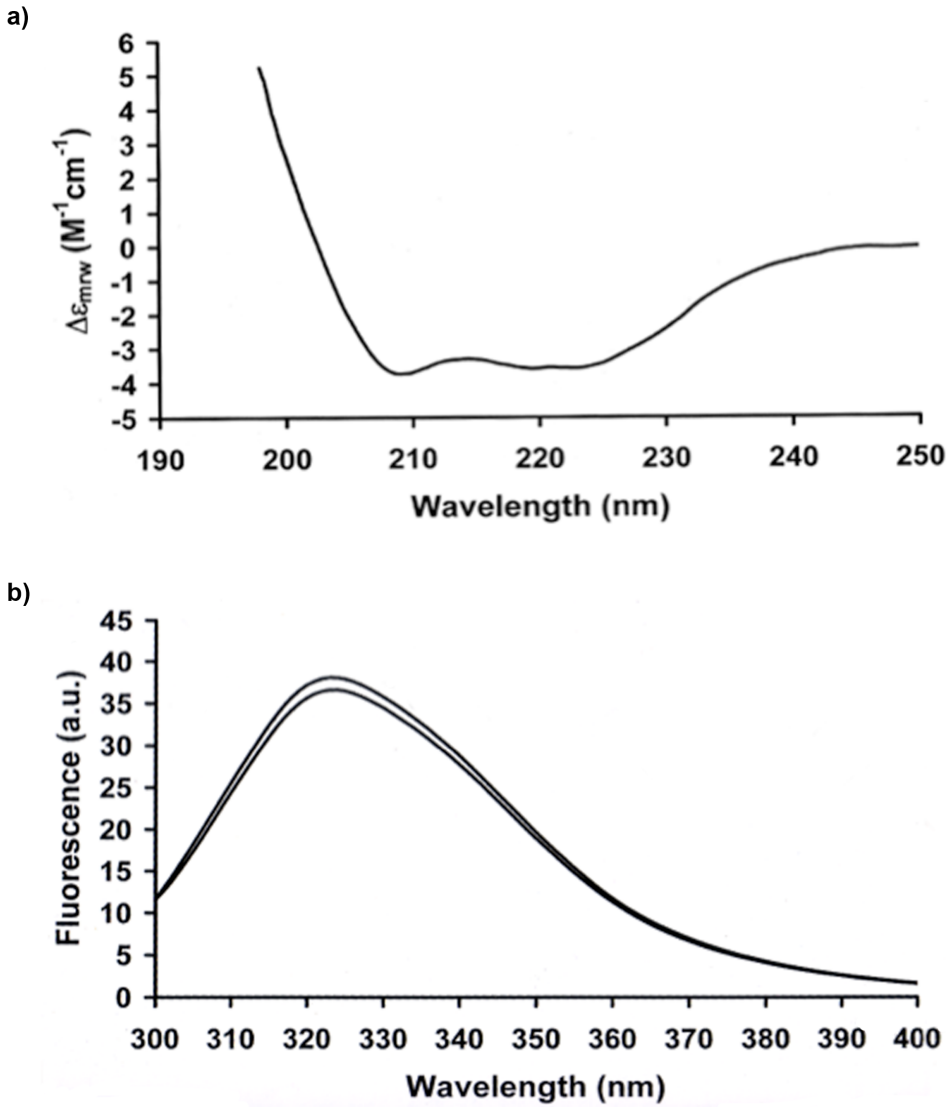


Figure 35: Measurements of CIR secondary structure.

a) A circular dichroism plot for PCHAS_000100, measured using a sample of 0.79 mg/ml concentration in PBS.

b) Comparison of PCHAS_000100 aromatic spectrum from different concentration samples: 0.79 mg/ml in PBS and 7.25 mg/ml in Tris-HCl containing 0.5 M proline.

4.4 Discussion

The expression of all three recombinant CIRs presented here describes the first soluble expression of PIR proteins. Expression of VIR proteins in *E. coli* has been described by Del Portillo and colleagues, who were able to express glutathione S-transferase (GST) fusion proteins from the second exon of two *vir* subfamily C genes (Fernandez-Becerra et al., 2005). This approach was extended to produce 22 VIR 'tags', which were differentially recognized by antibodies in the sera of *P. vivax* infected patients (Fernandez-Becerra et al., 2005). Their recognition by patient immune sera indicate that some determinants of specific Ab binding were still present in the recombinant VIR. However, since the authors noted that all GST fusion proteins were found in inclusion bodies during protein expression and were insoluble, it could be that the recombinant VIR were mis-folded and therefore only recognized by anti-VIR Abs when the proteins were denatured for SDS-PAGE and western blot.

In this study, a different expression system was chosen, *P. pastoris,* with the hope that the eukaryotic cellular machinery would directly produce correctly folded proteins. This expression system has been successfully used to express several merozoite proteins such as AMA1 (Kocken et al., 1999, Kocken et al., 2002), EBA-175 (Yadava and Ockenhouse, 2003) and MSP-1 (Brady et al., 2001, Hensmann et al., 2004) and the transmission-blocking mosquito-stage antigen *Pfs*25 (Zou et al., 2003).

Highly expressed *cir* genes representing different sub-families or with different characteristics of the *cir* family were chosen for synthetic gene design (discussed in chapter 3). However, very few data were available to know which genes were consistently highly expressed in different *P. chabaudi* infections and life-cycle stages. The genes used were chosen on the basis of preliminary measurements of *cir* transcription, using RT-qPCR and / or RNA sequencing data from a total of five *P. chabaudi* infected mice. Expression of recombinant CIRs to represent all of the sub-families and the most highly transcribed *cir*s would have been the optimal situation, in order to create a more representative library of CIRs. Since the assembly and annotation of the *P. chabaudi* AS genome sequence improved a great deal during this study, some early synthetic *cir* gene constructs were shelved as these no longer represented the real *cir* genes (data not shown). Under these experimental and time constraints, the expression of three recombinant CIRs was achieved, which could then be used to

investigate sub-group specific differences between the CIRs during *P. chabaudi* infection.

Modifications were incorporated into synthetic *cir* genes to facilitate expression of the recombinant proteins; these included the removal of predicted N-glycosylation sites and re-codonization for optimal transcription and translation in *P. pastoris.* Of the three proteins described here: PCHAS_000100, could be expressed in amounts up to 5 mg from 100 ml culture; whilst expression of the other two, PCHAS_070130 and PCHAS_040110, yielded relatively smaller amounts of protein, less than 0.5 mg from 100 ml culture. One possibility is that the low expression levels of both PCHAS_070130 and PCHAS_040110 was related to the inclusion of a TEV cleavage site to remove the six-his tag, which may have contributed to their poor protein yield (Kurz et al., 2006). This feature was not present in the design of PCHAS_000100. To address this in the future, the TEV cleavage site could be removed from the PCHAS_070130 and PCHAS_040110 constructs.

Another explanation for the low yield of recombinant CIRs PCHAS_070130 and PCHAS_040110 is that the pPIC9k vector used here required antibiotic selection after electroporation, in order to select for recombinant clones. Since the vector was linearized before electroporation, multimers could form, which may then have been inserted into the yeast genome within the 5' region of the AOX1 gene. Therefore increased antibiotic resistance was assumed to correlate with the number of plasmids acquired, and clones containing a higher copy number should have expressed the highest yields of the recombinant protein. However this may not always be the case, and highly resistant clones could have provided a false positive, where for example, resistance was acquired in the absence of multiple plasmid incorporation. In such instances, the clones that were highly antibiotic resistant would not correspondingly be the highest protein producers, which may have been the case for PCHAS_070130 and PCHAS_040110.

Future plans to select recombinant *P. pastoris* clones should focus on the initial selection of clones containing high numbers of integrated constructs to facilitate the expression of high protein yields. Initial attempts to lyse the yeast cells were unsuccessful due to poor activity of the lyticase enzyme used. Further attempts should be made to lyse the cells, following which the size of the integrated plasmids could be determined by PCR amplification using primers that spanned the plasmid and insertion sites, followed by gel electrophoresis.

Since cloning the recombinant *cir* genes, a new *P. pastoris* vector 'PINK-HC' has been released (Invitrogen), which allows less subjective recombinant clone selection due to adenine auxotrophy complementation. Using this system may increase the chance of initially selecting a clone with high protein expression, and reduce lengthy protein optimization investigations. Another advantage of this system is that, if proteins fail to secrete into the culture medium, there is a choice of seven signal sequences besides the traditional alpha factor (present in pPIC9k) that can be easily assayed to find which aids secretion of a particular protein.

Alternatively, a range of other eukaryotic expression systems exist, which may provide advantages over *P. pastoris* in the future production of recombinant CIR proteins; these have been used to express a range of *Plasmodium* proteins, illustrated in Table 21.

Table 21: Eukaryotic systems available for expression of recombinant proteins

Expression system:	Example of soluble antigen expression:	Efficacy of *Plasmodium* protein expression:	Major advantage compared to other systems:
Baculovirus-infected insect cells	The DBL3-X or the DBL5-ε domains of VAR2CSA (Mehlin et al., 2006)	41% of 17 proteins insoluble in *E. coli* were produced and soluble by insect cells expression (Chang et al., 1992)	Proteins are more likely to be folded than by yeast or *E. coli* expression. Yeast-expressed $MSP1_{19}$ was immunologically inactive; which was not true of $MSP1_{19}$ from this system (Lau et al., 2010)
Arabidopsis Thaliana	*Pf*$MSP1_{42}$ (Ghosh et al., 2002)	No large-scale analyses described to date.	May potentially be used as 'bioreactors' for production of high protein yields for vaccination (Srivastava et al., 2010)
Mammalian cells	The full-length extracellular region of var2CSA (Crosnier et al., 2010)	The extracellular domains of 49 *P. falciparum* proteins were expressed (including EBA-175 and MSP1 (Tsuboi et al., 2008)	Proteins can be targeted to sub-cellular compartments and the cell surface. Often used for investigation of ligand binding to host cells.
Wheatgerm cell free system	*Pf*s25, *Pf*CSP, and *Pf*AMA1 (Tsuboi et al., 2008)	65% of 93 expressed *P. falciparum* proteins were soluble using this system (Tsuboi et al., 2008)	There is no detriment of toxic products on host cell growth and proteins may be selectively labelled during synthesis.

Also listed in Table 21 is an expression system based on eukaryotic wheatgerm cell extract, this 'cell-free' system has been recently enhanced by Vinarov and colleagues to facilitate the expression of larger amounts of protein than previously possible (Vinarov et al., 2006). These different protein expression systems each have different advantages,

and which to use for future expression of CIR proteins would depend on the intended applications. For example, structural studies may require the incorporation of labels such as seleno-methionine to determine the protein fold, which would be facilitated by production of recombinant proteins using the wheatgerm cell free system.

PCHAS_070130 was predicted to be stable by ProtParam, meaning that the protein structure should remain constant. However, low expression levels and a tendency to form large aggregates made many studies using this protein unfeasible. In order to increase the yield of PCHAS_070130 expression and reduce protein aggregation, it may be necessary to remove the 80 N-terminal amino acids from the synthetic CIR, which contained the most predicted hydrophobicity. Such modification may prevent the subsequent recombinant protein from representing its native CIR, and thus may limit the utility of this protein for analyses of CIR function and Ab recognition.

PCHAS_000100 did initially consist of monomers in solution, which indicated that the protein was suitable for analyses of CIR structure. PCHAS_000100 secondary structure was observed to be 44% alpha helical by circular dichroism, suggesting that the recombinant protein was folded similarly to native CIRs, as unfolded proteins often predominantly contain random coils. This measurement also falls within the range predicted by Janssen and colleagues, that rodent PIR proteins would contain between 28 and 53 % alpha helices, suggesting that indeed this protein was folded like native CIRs (Janssen et al., 2004). Interestingly, the aromatic spectrum of PCHAS_000100 indicated that the tryptophan residue in this protein was not exposed to the buffer containing the protein sample. This residue is in fact found within the conserved amino acid motif 3 identified in chapter 2.3.3, present in the majority of CIR sequences. Such sequence conservation may arise from structural requirements, and the detection of one of these residues in the hydrophobic core of the protein indicates that this motif may be involved in maintenance of CIR protein structure.

Unfortunately in higher concentration preparations PCHAS_000100 had a tendency to aggregate, supporting the prediction by ProtParam (Gasteiger et al., 2005) that this protein was unstable. PCHAS_000100 was desalted into several different solutions, including Tris HCl pH 7.5 and Tris HCl containing 500 mM proline, in an effort to increase protein concentration without the formation of aggregates, but this was in vain. Due to the presence of monomers visible by SDS-PAGE (even when non-reduced, data not shown), it was hoped that regular crystals might be produced, that would be suitable for X-ray crystallography. However, in an initial trial of many different conditions for

crystal formation attempted by Dr L. Haire (Division of Structural Biology, NIMR), the only crystals observed were salt, so this avenue was abandoned.

To increase the chances of successful crystallisation, the tendency of the recombinant CIR proteins to aggregate must be addressed. This would probably include the removal of hydrophobic regions of the proteins, and regions that are predicted to self-complement, likely by drastically decreasing the size of the CIR protein. In this case it would be important to consider which regions of the protein would be important for correct folding, such as the conserved motif 3 that contains a tryptophan residue deeply buried in the protein core. It is likely that the highly conserved nature of this region is a consequence of functional constraints on protein folding, and the ability to generate structural information on the CIR protein core may provide valuable insights into the function of CIR proteins.

Chapter 5: CIR localization in infected red blood cells

5.1 Introduction

In order to determine the function of CIR proteins it is necessary to be able to identify native CIRs and to determine their location within the parasite and/or the iRBC. The first step to characterization of CIR proteins is to generate specific Abs, which can then be used to determine the size of native proteins by Western blotting, and their cellular location by immuno-fluorescence studies.

Polyclonal or monoclonal Abs can be generated using either the predicted extracellular domain of PIR proteins or to short peptides identified from the sequence data. In addition, recombinant proteins or peptides can be used to purify PIR-specific Abs induced by *Plasmodium* infection from the sera of immune individuals. These approaches have been used previously to generate specific Abs for the detection of native PIR, RIFIN, STEVOR and *Pf*EMP1 proteins (Del Portillo et al., 2001, Janssen et al., 2004, Cunningham et al., 2005, Fernandez-Becerra et al., 2005, Petter et al., 2007, Niang et al., 2009, Joergensen et al., 2010, Khattab and Meri, 2011).

During the initial description of the *vir* gene family, two GST-fusion VIR proteins were expressed in *E. coli*, and used to screen sera taken from patients infected with *P. vivax* (Del Portillo et al., 2001). Only three of 12 patients' sera recognized either of these proteins. The expression of full-length PIR proteins has proved difficult, as discussed in chapter 4. As the two recombinant VIR proteins were insoluble, they may thus not have represented native VIR proteins, which could explain their poor recognition by Abs present in the sera of *P. vivax* patients. Serum from one patient was affinity purified against 'VIR-C1-29', belonging to the phylogenetic sub-family C, and used to determine the localization of VIR proteins within iRBCs (Del Portillo et al., 2001).

An alternative solution to expressing full-length recombinant proteins is to synthesize short peptides and use these for the generation of polyclonal Abs. Whilst these peptides may represent amino acid residues hidden in nature due to the protein conformation, they are much easier to obtain. PIR peptide specific polyclonal Abs raised in rabbit, guinea-pig and mouse have been used in studies of VIR, YIR and CIR: to determine the molecular weights of native PIR proteins and to locate them near to or at the surface of iRBCs (Cunningham et al., 2005, Janssen et al., 2004, Fernandez-Becerra et al., 2005).

Janssen and colleagues generated Abs to the CIR peptide CSQKASEFVKSFKEL, found approximately in the centre of the CIR amino acid sequence (Janssen et al., 2004). Although this peptide was described as semi-conserved, only one CIR, PCHAS_000020, contained greater than 50% identity to the peptide by BLAST search, according to current annotation (http://www.genedb.org/blast/submitblast/GeneDB_Pchabaudi March 2011). These Abs were shown to recognize a protein of the expected size ~30kDa by western blot, and used to demonstrate that some CIR proteins appear to be located near the surface of fixed *P. chabaudi* trophozoite stage iRBCs (Janssen et al., 2004). The exact localization was uncertain because cellular fixation can permeabilize the iRBC membrane, allowing the entry of Abs. Thus, studies using fixed iRBC alone for protein localization may not distinguish proteins that are exposed at the iRBC surface from those which are located on the cytosolic face of the iRBC membrane.

Immunofluorescence and flow cytometric analyses have been carried out on unfixed and fixed mouse RBC infected with *Plasmodium yoelii* (Cunningham et al., 2005) using anti-sera generated to the semi-conserved YIR peptides KLYDALQSLCNMYNEF and ISAGCLYLLDEFIKDC (found in 50% and 21% of YIRs, respectively). This demonstrated the presence of YIR on the surface of schizont iRBC as well as intra-cellularly on fixed and permeabilized iRBC (Cunningham et al., 2005).

The objectives of the work described in this chapter were to:

- **i)** Generate CIR-specific Abs using a peptide conserved within the CIR repertoire, and to use these Abs for:
- **ii)** Detection of CIRs expressed within iRBCs by SDS-PAGE and western blotting.
- **iii)** Determination of where and when CIR proteins were expressed in iRBCs using immunofluorescence confocal microscopy and flow cytometry.

5.2 Methods

5.2.1 Preparation of *P. chabaudi* infected material

Origins and housing of parasites and mice have been described previously, chapter 3.2.4. Parasites expanded from frozen stabilate were passaged twice through BALB/c mice kept under normal or reverse light conditions to ensure maximal synchronicity. Late-ring and early-trophozoite iRBC stages were obtained from *P. chabaudi* infected mice housed under normal light conditions (light between 7 am and 7 pm), whilst all other stages of parasite development were obtained from *P. chabaudi* infected mice housed under reverse light conditions (light between 7 pm and 7 am), where schizogony occurred between 12pm and 1pm daily.

Blood samples were taken from mice either under terminal anasthaesia, for collection of the total blood volume, or by removal of the tail tip, for samples less than 100 µl. Blood samples were taken into Krebs saline (128 mM NaCl, 4 mM KCl, 2 mM $CaCl_2$, 1 mM $MgSO_4$, 1 mM NaH_2PO_4, 25 mM $NaHCO_3$, (Krebs and Eggleston, 1940) containing 0.2% glucose and 25 U/ml heparin (Leo Pharmaceuticals).

Saponin lysis of RBCs was performed in order to remove the abundant haemoglobin present within the cells, which may be recognized non-specifically by Abs during western blotting. Lysis was carried out using naïve or *P. chabaudi* infected blood as follows: Blood was first washed three times in PBS, using low-speed centrifugation at 2000 x *g* for 5 minutes at 4 °C to pellet RBCs. Cells were then suspended in 0.15% saponin (Sigma), in PBS, for 10 minutes, followed by centrifugation at 2000 x *g* for 5 minutes at 4 °C. The lyzed cellular fraction was found in the pellet from this step, hereafter referred to as the 'parasite fraction'.

After removal of the parasite fraction, the supernatant was then subjected to ultra-centrifugation at 100,000 x *g* using the fixed angle rotor TLA 100.3 (Beckman) for 30 minutes at 4 °C to pellet the lysed membranes, the 'membrane fraction'. Both the parasite and membrane fractions were washed three times by re-suspension in PBS and further centrifugation at either 2000 x *g* for 5 minutes at 4 °C or 100,000 x *g* for 30 minutes at 4 °C, respectively.

Infected RBC membrane fractions are comprised both of the iRBC surface and PV membranes, as these are derived from the RBC membrane (Siddiqui et al., 1979). The

parasites' own membranes are resistant to saponin lysis, at least in *P. falciparum* (Siddiqui et al., 1979). PV and iRBC membranes were used directly for electrophoresis.

Membranes of uninfected RBCs were either used directly for electrophoresis or to absorb-out non-specific recognition of RBC surface proteins from the anti-1332 serum.

5.2.2 CIR peptide design and synthesis

A conserved amino acid motif identified within the CIR repertoire, chapter 2.3.3, was used as the basis for design of the conserved CIR peptide. The sequence AEYAILWLCYKIN, derived from the amino acid motif, was detected by BLAST in 15.82% of CIR sequences, allowing 0 - 2 mis-matches. An N-terminal cysteine residue was added to facilitate peptide conjugation to the carrier proteins keyhole limpet haemocyanin (KLH) and bovine serum albumin (BSA). Peptide synthesis and conjugation of 5 mg peptide to each carrier protein was carried out by Jerini Peptide Technologies, Berlin, Germany.

5.2.3 Generation of anti-sera

Mouse anti-sera was prepared using 10 BALB/c mice, which were given three intra-peritoneal immunizations of 50 μg of the KLH conjugated conserved CIR peptide in the presence of Sigma adjuvant (Sigma), prepared according to manufacturer's instructions.

In addition, two New Zealand White rabbits were immunized three times sub-cutaneously with 250 μg of the KLH conjugated conserved CIR peptide in the presence of Titermax ® adjuvant (CytRx Corporation), carried out by Harlan laboratories, Loughborough, UK.

Sera from immune and naïve mice and rabbits were prepared by collection of blood in the absence of heparin, which was allowed to clot at room temperature for 2 hours. Low speed centrifugation was then carried out at 2000 x *g* for 10 minutes to remove the clot, and supernatant was subjected to high speed centrifugation 13,000 x *g* for 5 minutes to remove aggregated material, debris and free erythocytes.

5.2.4 Detection of CIR-specific Ab titres in anti-sera

96-well flat Polysorb plates (Nunc, Denmark) were coated with the BSA-conjugated CIR peptide at 5 μg/ml peptide concentration, diluted in PBS, and 50 μl was applied per well. Plates were incubated overnight at 4 ˚C, following which, the residual surface was blocked with 200 μl PBS containing 1% BSA, 0.3% Tween 20 and 0.05% Sodium azide (blocking buffer).

Plates were washed three times with PBS containing 0.9% sodium chloride, and 10% Tween 20, pH 7.2. Sera from the immunized mice and rabbits were compared before and after immunization. Naïve and immunized mouse sera were pre-diluted to 1/100, whilst the initial dilution for rabbit sera was 1/10,000. Serial two-fold dilutions were performed in blocking buffer. Negative controls containing only blocking buffer were treated in the same way. The primary antibody was incubated for 1 h at 37 ˚C, followed by three washes with PBS containing 10% Tween 20, as described above.

CIR-specific Abs were detected using alkaline phosphatase-conjugated goat anti-mouse Ig (Southern Biotechnology), diluted to 2 μg/ml, and incubated for 1h at 37 ˚C, followed by three washes in PBS alone. Bound antibodies were then visualised with 50 μg of 1 mg/ml p-nitropheny-phosphate (p-NPP) in diethanolamine buffer (0.26 M diethanolamine, 8 mM magnesium chloride, pH 9.8). The optical density (OD) was measured at the reference wavelength 405 nm.

The titres of specific Abs induced by immunization were calculated by interpolation of the titration curve, using non-linear regression analysis performed in Prism v5.0. Here, titre was defined as the dilution at which the pooled immunized sera reached the baseline OD 405 nm of the pooled sera prior to immunization.

5.2.5 Preparation of anti-sera

The anti-sera generated in mouse and rabbit were evaluated using western blots of naïve and *P. chabaudi* infected RBCs, to determine whether the Abs specifically recognized parasite proteins. The rabbit anti-serum was observed to contain higher proportions of Abs recognizing RBC proteins than the anti-sera generated in mice. For this reason, the sera were prepared differently for use in subsequent applications, using methods which had been found to remove RBC cross-reactivity.

Pre-immune sera and anti-sera to the conserved CIR peptide, generated in rabbit, were absorbed using RBC membrane fractions from uninfected mice, which were prepared as described above. 100 µl serum and 50 µg membrane fraction were incubated overnight, rotating, at 4 °C. To remove un-bound RBC membranes, samples were then subjected to ultra-centrifugation at 100,000 x *g* for thirty minutes at 4 °C.

A simpler procedure was carried out for pre-immune and anti conserved CIR peptide sera generated in mice, whereby absorption was carried out using intact RBC from uninfected mice. 100 µl serum and 50 µl packed RBC were incubated overnight, rotating, at 4 °C. To remove un-bound RBCs, the samples were then subjected to centrifugation at 2000 x *g* for 5 minutes at 4 °C. To have adequate material, sera generated from a pool of 10 conserved CIR peptide-immunized mice was used for further investigation of CIR proteins.

For simplicity, both RBC membrane-absorbed anti-sera generated in rabbit and RBC-absorbed anti-sera generated in mice will be referred to as anti-CIR sera.

5.2.6 SDS PAGE and immuno-blotting

Proteins were resolved using 15% Tris-glycine gels, prepared according to Table 22, in running buffer containing 0.5% SDS (v/v, Invitrogen), 0.126 M Tris base (Sigma) and 1 M Glycine (BDH Biosciences), at 200 volts for 40-60 minutes. Reducing conditions were used throughout, where samples contained 100 mM DTT. Proteins were run alongside 1x SeeBlue Plus2-prestained standard (Invitrogen).

Table 22: Composition of 15% Tris glycine SDS-PAGE gels.

Reagents	Resolving gel	Stacking gel
Resolving buffer (1.5 M Tris, Invitrogen, pH 8.8)*	1.24 ml	-
Stacking buffer (0.5 M Tris, Invitrogen, pH 6.8)*	-	0.75 ml
Distilled water	1.24 ml	1.776 ml
30% Acrylamide/Bis solution (29:1 ratio, BioRad)	2.48 ml	0.474 ml
10% Ammonium persulfate (APS, Sigma-Aldrich)	0.03 ml	0.012 ml
Tetramethylethylenediamine (TEMED, Sigma)	0.33 µl	3.9 µl
Total composition of a single gel:	**5 ml**	**3 ml**

* Buffers prepared using Tris base (Sigma) and containing 0.4% SDS (v/v from 20% stock, Invitrogen)

Total protein concentration in each sample was determined by BCA assay (Pierce), as described in chapter 4.2.2. For analysis of *P. chabaudi* infected- or uninfected RBC-derived material, 20 μg protein was loaded onto the SDS-PAGE gels. The recombinant CIR proteins PCHAS_000100 and PCHAS_040110, described in chapter 4, were compared with recombinant $MSP1_{p21}$ (Hensmann et al., 2004), which was generated in the same expression system as the recombinant CIRs, *P. pastoris*, and contained an identical C-terminal six-histidine tag. For analysis of each recombinant protein, 500 ng of each was loaded onto the SDS-PAGE gels.

All gels were used for immuno-blotting, where the proteins were electrophoretically transferred to Hybond C membrane (Amersham Biosciences) in buffer containing 20% methanol (v/v), 0.025 M Tris base (Sigma) and 0.2 M Glycine (BDH Biosciences) for 3 h at 30 V. Membranes were blocked at 4 ˚C overnight in PBS containing 0.1% Tween 20 (Sigma) and 3% bovine serum albumin (BSA), and washed for 10 minutes three times with PBS containing 0.01% Tween 20.

Specific proteins were detected using a variety of Abs, diluted in PBS containing 0.1% Tween 20 and 3% BSA. *P. chabaudi* derived material was probed with the conserved CIR peptide rabbit anti-serum at 1:500 dilution, or a combination of the monoclonal Abs Ter119 and NIMP23, which recognize Glycophorin A and $MSP1_{p21}$ respectively (Kina et al., 2000, Boyle et al., 1982, McKean et al., 1993), used at concentrations of 32 pg/ml and 18 μg/ml. In addition, recombinant CIR were detected using conserved CIR peptide rabbit anti-serum, normal serum (NS) from the pre-immune rabbit, both at 1:500 dilution; or an antibody which recognizes the six-histidine tag of the recombinant proteins, used at 0.5 μg/ml (Novagen).

Bound antibodies were detected using secondary horseradish peroxidase (HRP)-conjugated Abs: goat anti-rabbit IgG, goat anti-mouse IgG and goat anti-rat igG (Biorad) were used at 1:10,000 dilution. Membranes were washed as before, upon which SuperSignal West-Pico chemi-luminescent substrate (Thermo Scientific) was added, according to manufacturer's instructions, and incubated in the dark for 5 minutes. Light emission was detected by exposure to Biomax film (Kodak).

5.2.7 Immuno-fluorescence assays

P. chabaudi AS infected blood was collected from BALB/c mice on day 7 of infection, directly into Krebs saline (114 mM NaCl, 4.57 mM KCl, 1.15 mM $MgSO_4$(Krebs and

Eggleston, 1940), containing 0.2% glucose and 25 U/ml heparin (Leo Pharmaceuticals). The iRBCs were centrifuged at 1500 x *g* at 4 ˚C for 5 minutes, and re-suspended in foetal calf serum. This material was used to make thin blood films on glass slides, which were allowed to dry, and then fixed using a solution containing 90% acetone and 10% methanol, then dried. Slides were then incubated in PBS containing 3% bovine serum albumin (Sigma). All Abs were diluted in PBS containing 3% BSA, and 100 µl applied directly to the fixed iRBC on each slide, underneath a cover-slip. Primary reagents were rabbit anti-CIR serum (or pre-immune serum) diluted 1/50, NIMP23 at a concentration of 50 µg/ml and Ter119 conjugated to biotin (Biolegend) at 50 µg/ml. Secondary reagents were goat anti-rabbit IgG conjugated to Alexa 680, goat anti-mouse IgG conjugated to Alexa 594 and Streptavidin conjugated to Alexa 750-Allophycocyanin (all Molecular Probes, used at 10 µg/ml). Incubations were carried out at 37 ˚C for one hour or 4 ˚C overnight in a humid chamber. Abs were applied sequentially, to prevent non-specific recognition of other Abs by the polyclonal antisera, in the following order: rabbit anti CIR Abs, goat anti-rabbit Alexa 680, followed by NIMP23 and Ter119 and their secondary reagents.

Between each Ab incubation three washes were performed, by placing the slide into 50 ml PBS for five minutes. Parasite nuclei were visualised using 1 mg/ml 4',6-Diamidino-2-phenylindole (DAPI, Sigma Aldrich) in PBS. Slides were mounted in anti-fadent AFI™ (Citifluor), sealed using nail polish and stored at 4 ˚C in the dark. Immunofluorescence of iRBC was visualised using the Leica SP2 confocal microscope with total magnification of 2500 x, 405 nm, 488 nm, 594 nm and 750 nm lasers and Leica software. As the aim was to localize the signal of anti-CIR staining, the settings for gain and offset were determined empirically for each slide.

5.2.8 Schizont culture

P. chabaudi AS infected blood was collected from BALB/c mice on day 7 of infection, directly into Krebs saline (114 mM NaCl, 4.57 mM KCl, 1.15 mM $MgSO_4$, (Krebs and Eggleston, 1940) containing 0.2% glucose and 25 U/ml heparin (Leo Pharmaceuticals). Blood was diluted 1:4 using RPMI 1640 medium (Gibco, to which 6 mM HEPES and 2 mM L-glutamine (both Invitrogen) were added), pre-warmed to 37 ^{o}C. Diluted blood was immediately washed twice using 50 ml, pre-warmed to 37 ^{o}C, complete RPMI 1640 medium [Gibco (containing 6 mM HEPES, 0.5 mM sodium pyruvate, 2 mM L-

glutamine, 50 μM β-Mercapto-ethanol (Invitrogen) and 10% heat inactivated foetal calf serum (PAA laboratories, Pasching, Austria)], and followed by centrifugation at 600 x *g* for 5 minutes at 37 ^{o}C.

Infected RBCs were then transferred to non-vented culture flasks (Becton Dickenson) containing 5 ml complete RPMI medium, pre-warmed to 37 ^{o}C, for every millilitre of iRBC. The gas composition was adjusted to 10% O_2, 5% CO_2, 85% N_2 and iRBCs were incubated at 37 ^{o}C for between 3 and 4 hours until schizonts were observed to predominate in the culture and/or free merozoites started to appear.

5.2.9 Flow cytometry

P. chabaudi AS infected blood was collected from BALB/c mice on day 7 of infection, directly into Krebs saline (114 mM NaCl, 4.57 mM KCl, 1.15 mM $MgSO_4$, (Krebs and Eggleston, 1940) containing 0.2% glucose and 25 U/ml heparin (Leo Pharmaceuticals). *P. chabaudi* AS iRBC were filtered through Plasmodipur filters (EuroProxima) to remove leukocytes and then washed three times in PBS, using centrifugation at 1500 x*g* at 4 ˚C for 5 minutes. Red blood cells were enumerated using a haemocytometer (NeuBauer), and blood diluted accordingly to give a final concentration of $1x10^7$ RBCs per ml in PBS containing 1% BSA (w/v), 2 mM ethylene-diamine-tetra-acetic acid (EDTA) and 0.01% sodium azide (FACS buffer).

The RBCs ($1x10^6$) were plated in triplicate into V-bottom plates (Nunc), centrifuged at 1200 x *g* for 1 minute at 4 ˚C, and the supernatant removed. Cells were first incubated with mouse pre-immune sera or conserved CIR peptide anti-sera diluted 1/50, followed by goat anti-mouse IgG conjugated to FITC (Southern Biotechnology) used at 10 μg/ml concentration. All Abs were diluted in FACS buffer, and incubations performed in a final volume of 100 μl on ice for 30 minutes. Once fluorescently labelled Abs were added, all steps were performed in the dark. After all antibodies had been added, the cell suspensions were incubated with 10 μg/ml Hoechst 33342 (Invitrogen), prepared in FACS buffer, for 10 minutes on ice, to detect parasite DNA. Between each incubation step, cells were re-suspended twice in 150 μl FACS buffer, followed by centrifugation at 1200 x *g* for 1 minute at 4 ˚C.

Single colour and unstained controls were included for all samples (including uninfected RBCs, data not shown). Immediately prior to data acquisition, all samples were passed through a 0.45 μm filter (Nunc), to ensure no cellular aggregates were

present. A minimum of 300,000 events were acquired using the LSR II flow cytometer (BD biosciences), using default filter settings, with FACSDIVA software (BD). Following data collection, information was further analyzed using FlowJo software version 8.8.6 (Tree Star).

5.3 Results

5.3.1 Preparation of CIR-specific polyclonal Abs

In order to detect native CIR proteins within *P. chabaudi* AS iRBCs, reagents were required that would recognize a large proportion of the CIR repertoire. The amino acid motif, identified as 'motif 3' in chapter 2.3.3, was present in the majority of CIR sequences and within the predicted extra-cellular domain of the proteins. In particular, the residues YAILWL within the amino acid motif were present in over 75% of CIRs. This is shown as a weblogo image (Crooks et al., 2004) in Figure 36a, where the proportion of CIRs containing each residue of the amino acid motif is depicted by the height of the letter. A consensus sequence was derived from this amino acid motif, Figure 36b, which was present in 64% of CIRs by basic local alignment tool analysis [BLAST, (Altschul et al., 1990)], allowing up to five mis-matches. A peptide was synthesized based upon the sequence AEYAILWLCYKIN, to which an N-terminal cysteine residue was added to facilitate conjugation of the peptide to the carrier proteins keyhole limpet haemocyanin (KLH) and bovine serum albumin (BSA).

The KLH-conjugated peptide was used to immunize both rabbits and mice, for the generation of CIR-specific polyclonal Abs. Enzyme linked immunosorbent assays (ELISAs) were performed using the antisera to calculate the titre of Abs recognizing the peptide. Rabbit sera contained higher titres of anti-CIR Abs, with a mean titre of 5.12×10^6 (+/- SEM of 1.36, calculated from triplicate titrations), than mouse sera, which had a mean titre of 1.04×10^5 (+/- SEM of 0.044, calculated from triplicate titrations). For this reason, the rabbit polyclonal anti-sera were preferentially used for detection of CIR proteins.

Rabbit anti-CIR Abs were investigated via western blot, to confirm whether they exclusively recognized CIR proteins or could recognize other *P. chabaudi* proteins non-specifically. Initial western blots using the polyclonal Abs resulted in the recognition of uninfected RBCs in addition to *P. chabaudi* iRBCs. To reduce this, the anti-serum was first absorbed on intact RBC, which did not entirely remove the observed cross-reactivity for the Abs generated in rabbit. Subsequently the anti-1332 serum was absorbed against RBC membranes. This resulted in recognition of iRBC by the anti-CIR Abs, without recognition of uninfected RBC, shown in **Figure 37**a. The *P. chabaudi* proteins recognized by the anti-CIR Abs were approximately 45, 35, 25 and 20 kDA, within the size range predicted for native CIR proteins, based upon current *cir*

gene annotation. The majority of CIR proteins contain between 311 and 490 amino acids, which approximately equates to predicted molecular weights of between 34 and 54 kDa, whilst the smallest predicted CIRs would have molecular weights ranging upwards from 15 kDa.

In parallel, the same material was probed with a mixture of Abs (**Figure 37**b), as a loading control to ensure that equivalent amounts of infected and uninfected RBC membrane preparations were loaded onto the SDS-PAGE gel. The rat monoclonal Ab Ter119, which recognizes a Glycophorin A-associated protein, not glycophorin A itself on the RBC membrane (Kina et al., 2000), detected bands of approximately 30, 50 and 64 kDa in both infected and uninfected RBC. These correspond with proteins of 32, 52 and 60 kDa previously described for Ter119 (Boyle et al., 1982, McKean et al., 1993). In addition, the monoclonal Ab NIMP23, which recognizes the 21 kDa fragment of *P. chabaudi* MSP1 [the equivalent of *P. falciparum* $MSP1_{19}$, (Kina et al., 2000, Boyle et al., 1982, McKean et al., 1993)], detected MSP1 only in iRBCs. This indicates that the uninfected RBC material was not contaminated with parasites, as no MSP1 was detected.

Figure 37b confirms that the sizes of proteins recognized by the anti-CIR Abs were not the same as the sizes of predominant parasite and RBC proteins. However, this did not necessarily indicate that the proteins recognized by anti-CIR Abs were indeed CIRs. To confirm that anti-CIR Abs recognized CIR proteins, western blots were carried out using two of the recombinant CIR described in chapter 4, PCHAS_040110 and PCHAS_000100, **Figure 37**c demonstrates that the anti-CIR serum recognized CIR proteins specifically, with no recognition of another recombinant *P. chabaudi* protein, $MSP1_{p21}$, produced using the same expression system, *P. pastoris* (Hensmann et al., 2004). Antibodies to the His-tag, present on all recombinant proteins, confirmed that equal amounts of protein were loaded on the gel, indicating that the failure of anti-CIR serum to recognize MSP1 was not due to different protein concentrations, but rather that it specifically recognized CIR proteins.

In summary, the anti-CIR Abs generated in both mouse and rabbit were able to recognize proteins of the expected range of molecular weights for native CIR in *P. chabaudi* iRBC, by western blot. Importantly, these Abs also recognized recombinant CIR proteins from both of the major CIR subfamilies identified in chapter 2, **Figure 37**c. Therefore these Abs appeared to be CIR-specific and able to detect a wide range of CIR proteins.

5.3.2 Localization of CIR proteins in *P. chabaudi* iRBCs

To inform which stages of parasite development should be analyzed for the cellular localization of CIR proteins within iRBCs, blood samples were taken every two hours during the erythrocytic development cycle of *P. chabaudi*, which grows synchronously *in vivo* (Hawking et al., 1972). A representative Giemsa-stained iRBC from each time in the intra-erythrocytic development of *P. chabaudi* is depicted in Figure 38a. From these results, time-points corresponding to early- and late- phases of ring stage and trophozoite stage iRBCs were used for immunofluorescent analysis of CIR localization (Figure 38b). Infected RBCs were co-stained with Abs recognizing $MSP1_{p21}$ and the iRBC surface, to allow some determination of CIR positioning. Controls ensuring that the Abs used did not bind to iRBCs non-specifically or cross-react with other reagents are attached in Appendix 5.1.

In ring- and early trophozoite- iRBC stages, the pattern of immuno-fluorescence was similar. The parasitophorous vacuole (PV) around developing parasites could be clearly seen by $MSP1_{p21}$ staining [as described by Smythe and colleagues (Smythe et al., 1988)], Figure 38b. Anti-CIR staining remained close to the parasite nuclei in all iRBCs, in a similar but more diffuse pattern than $MSP1_{p21}$. This indicates that CIR proteins were located within the PV membrane, and diffuse staining also occurred beyond the PV membrane in some ring- and early trophozoite- stage iRBCs.

In late stage trophozoite iRBCs, CIRs were still detected close to the parasite nuclei, but in more than 50% of cells (of 25 counted), were also observed in the iRBC cytoplasm, Figure 38b. A small number of iRBCs were enumerated as cytoplasmic localization of CIRs was only clearly visible by confocal microscopy. A magnified view of an example trophozoite stage iRBC with such peripheral CIR staining is depicted in Figure 39a. Computationally, an axis was drawn across the trophozoite and the fluorescence intensity of anti-CIR, anti-MSP1 and nuclear staining were determined along this line, allowing the detailed study of the relative position of cellular components. The fluorescence intensity related to MSP1 appeared to define the edges of the PV, whilst CIR fluorescence was localized around the parasite nucleus, and in a discrete peak close to the edge of the iRBC.

In addition, because the majority of *P. chabaudi* schizont stages sequester in the deep vasculature (Gilks et al., 1990, Mota et al., 2000), iRBCs at the late trophozoite stage

were cultured *in vitro* to reach schizogony and the release of free merozoites. These parasites were used for immunofluorescence analysis, shown in Figure 39b and c. Schizonts containing developing merozoites displayed the same CIR localization patterns as those found on free merozoites, which consisted of localization of CIR to the apical end of the merozoite. An axis was drawn computationally across a representative merozoite, as described above for the trophozoite stage iRBC. Here, the fluorescence intensity related to MSP1 defined the whole merozoite surface (Crooks et al., 2004). CIR fluorescence did not colocalize with MSP1, but rather appeared most intense at a discrete point at the apical end of the merozoite.

Together, the data presented in Figure 39 indicated that CIR proteins were expressed in all parasites, at the apical end of merozoite stages, and within the PV of ring- and trophozoite iRBC stages. The detection of CIR close to the iRBC membrane of at least half of late-trophozoite stage parasites indicated that some CIR family members could be exposed at the iRBC surface.

5.3.3 Detection of CIRs at the surface of live iRBCs

To investigate the possibility that CIRs were exposed at the surface of some iRBCs, live (un-permeabilized) iRBCs at the late trophozoite stage were stained with anti-CIR Abs prepared in mouse. Flow cytometry was then used to determine the proportion of iRBC containing surface-exposed CIR during *P. chabaudi* infection, Figure 40.

Data were analyzed using FlowJo software (Tree Star). Firstly, a region was drawn within the forward and side scatter parameters - to exclude very small, granular or other cells which were the wrong size or granularity for iRBC, Figure 40a. Subsequently, Hoechst 33342 was plotted against CIR fluorescence, and a region drawn to define the DNA positive iRBCs, Figure 40b. This definition may have also included some reticulocytes, as previously discussed in chapter 3. The percentage of CIR-positive iRBC is shown as a histogram in Figure 40c. Although the percentage of CIR positive iRBC was small, iRBC stained with pre-immune sera from the same mice were almost completely negative (0.01%). These data suggest that a small proportion of iRBCs expressed CIR at the surface, and confirm the localization of some CIRs at this position observed by immuno-fluoresence, described above.

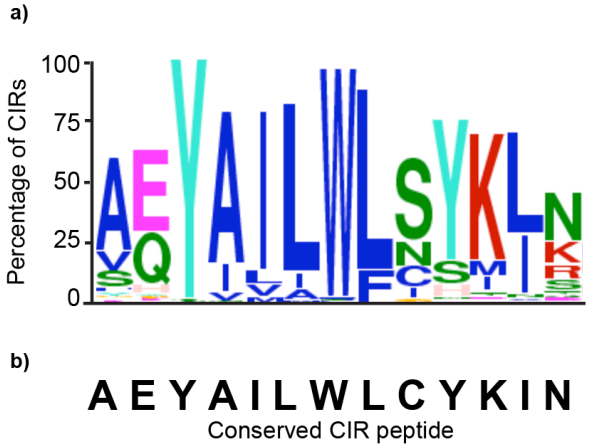


Figure 36: Design of the conserved CIR peptide

A highly conserved CIR amino acid motif was detected in CIR amino acid sequences, described in chapter 2 (Table 7).

- **a)** The weblogo image (Crooks et al., 2004) shows the proportion of all 198 CIR amino acid sequences containing each residue of this motif.
- **b)** A consensus sequence was taken from this motif, shown below the weblogo image. This sequence was synthesized as the conserved CIR peptide.

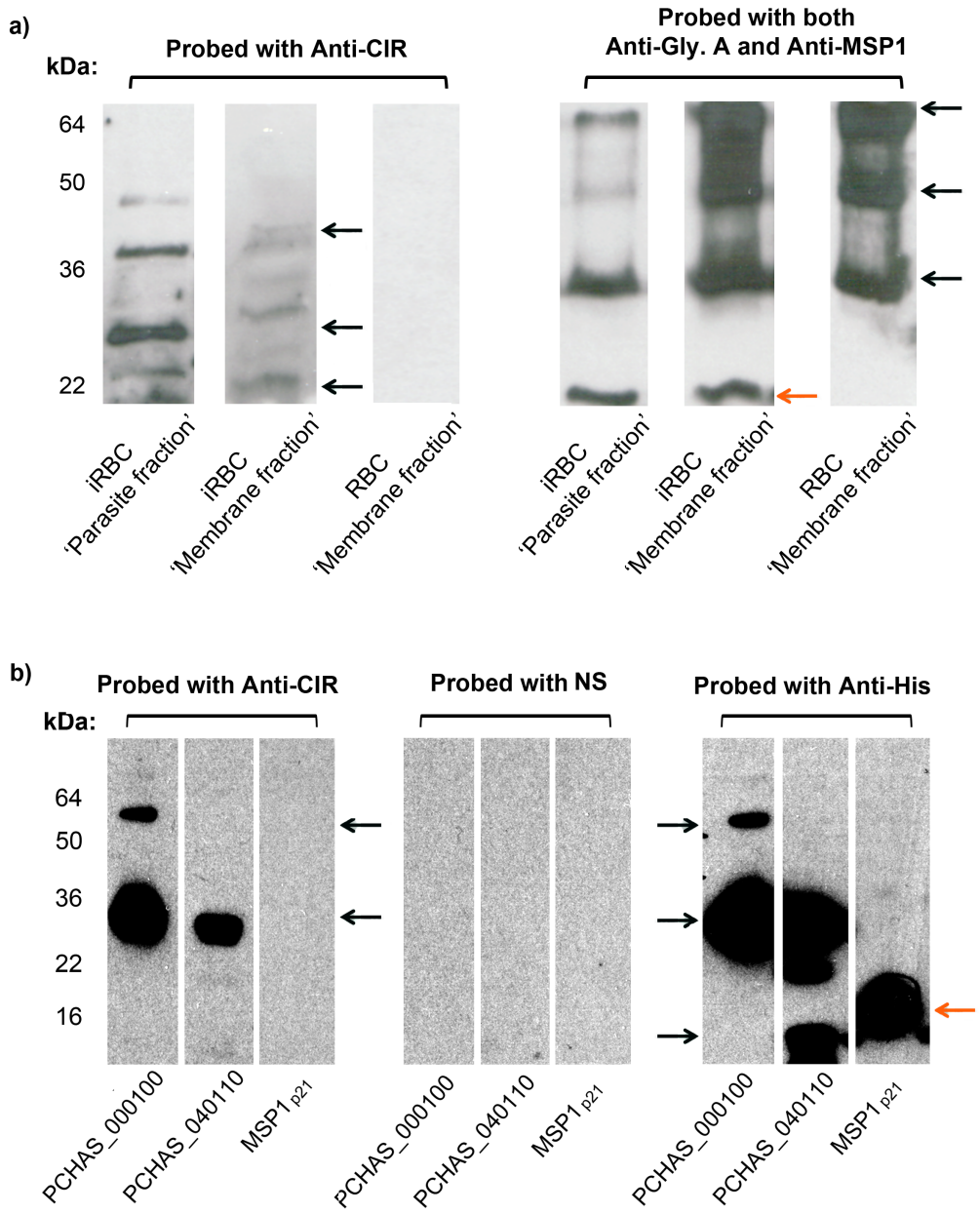


Figure 37: Specificity of polyclonal anti-CIR Abs.

a) 20 µg of protein was loaded (left to right) from: total protein and membrane fractions extracted from *P. chabaudi* iRBC, and the membrane fraction of uninfected RBC. These were resolved using 15% Tris-glycine SDS-PAGE gels and used for western blots, which were probed with polyclonal anti-CIR Abs, generated in rabbit (left panel). Black arrows indicate CIR-specific bands in the iRBCs. The same membrane was stripped and re-probed with a combination of Ter119 and NIMP23 (right panel), which recognize the RBC membrane and $MSP1_{p21}$, respectively (Kina et al., 2000, Boyle et al., 1982, McKean et al., 1993). Black arrows indicate bands recognized by Ter119, whilst a red arrow indicates bands recognized by NIMP23.

b) The recombinant CIR proteins PCHAS_000100, PCHAS_040110 and recombinant $MSP1_{p21}$ (Hensmann et al., 2004), were resolved using 15% Tris-glycine SDS-PAGE gels, onto which 500 ng of each recombinant protein was loaded. Western blots were carried out, probing the proteins with anti-CIR Abs, pre-immune serum from the same rabbit (NS), or an Ab recognizing the hexa-histidine tag present on all recombinant proteins (anti-His). Black arrows indicate CIRs, $MSP1_{p21}$ is indicated by a red arrow.

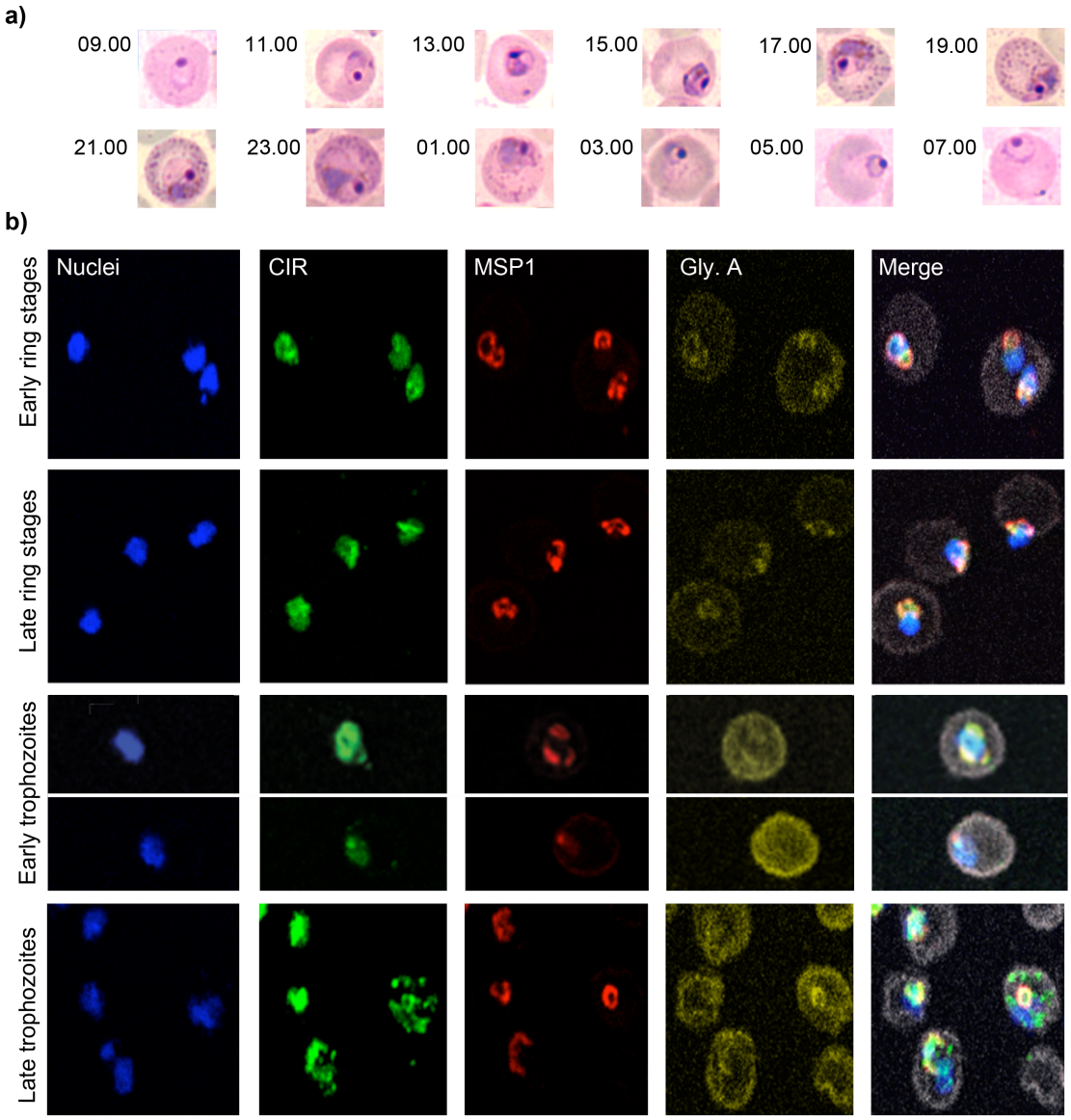


Figure 38: CIR localization by indirect immunofluorescence and confocal microscopy during the erythrocytic growth cycle of *P. chabaudi*.

a) A representative example of a Giemsa stained iRBC, taken every 2 hours during the 24h erythrocytic growth cycle, from 09.00 hours until 07.00 hours the following day.

b) Immunofluorescence analysis of specific stages of development: early ring (05.00 h), late ring (11.00 h), trophozoite (17.00 h) and late trophozoite (23.00 h).

P. chabaudi infected RBCs were probed with: DAPI, for detection of parasite nuclei; rabbit anti-serum to the conserved CIR peptide, detected using goat anti-rabbit Alexa 488; NIMP23, which recognizes MSP1, detected using goat anti-mouse Alexa 594 and Ter119, which recognizes the RBC membrane, detected using streptavidin Alexa 750-Allophycocyanin. These are shown from left to right, individually, followed by merged images (right).

Images of control- stained iRBCs are attached in Appendix 5.1

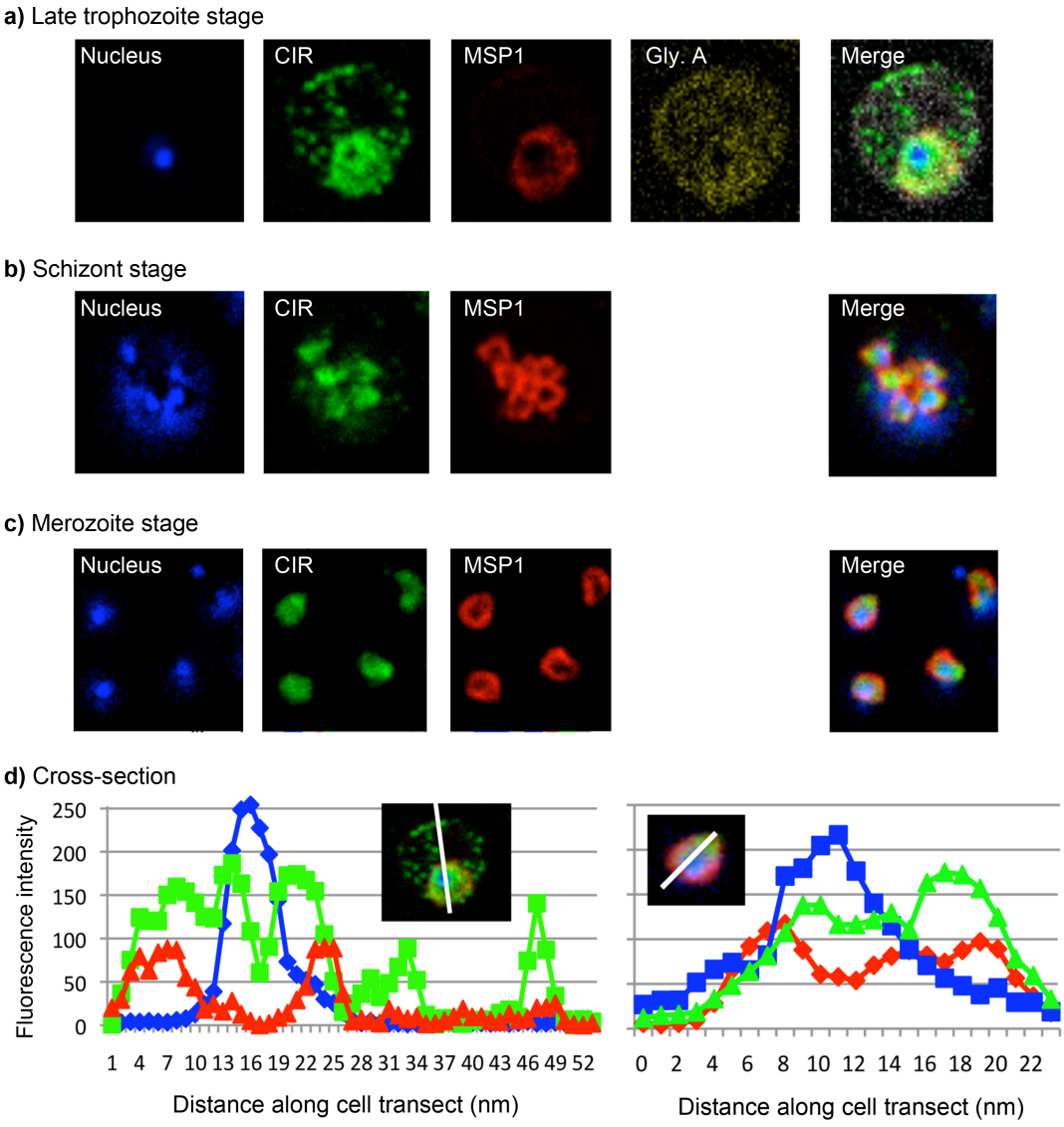


Figure 39: CIR localization within individual *P. chabaudi* trophozoite, schizont and merozoite stages of development.

Trophozoite stage iRBC, **a)**, taken at 11.30 am from *P. chabaudi* infected mice kept under reverse light conditions (in darkness from 7am until 7pm). Schizont stage iRBC, **b)**, and free merozoites, **c)**, after 4 h *ex vivo* culture of iRBCs

Infected RBCs were probed with: DAPI, for detection of parasite nuclei; rabbit anti-serum to the conserved CIR peptide, detected using goat anti-rabbit Alexa 488; NIMP23, which recognizes MSP1, detected using goat anti-mouse Alexa 594 and Ter119, which recognizes the RBC membrane, detected using streptavidin Alexa 750-Allophycocyanin. These are shown from left to right, followed by merged images.

Graphs showing the intensity of fluorescence across an individual trophozoite stage iRBC and merozoite were produced, **d)** [using the merged images shown in **a)** and **c)**]. "Plot profiles" were created using Image J software (Abramoff et al., 2004), along the specified cellular axis shown with a white line. In these graphs, the fluorescence intensity of DAPI was represented with blue lines, the fluorescence intensity associated with MSP1-specific staining was represented with red lines and the fluorescence intensity associated with CIR-specific staining was represented with green lines.

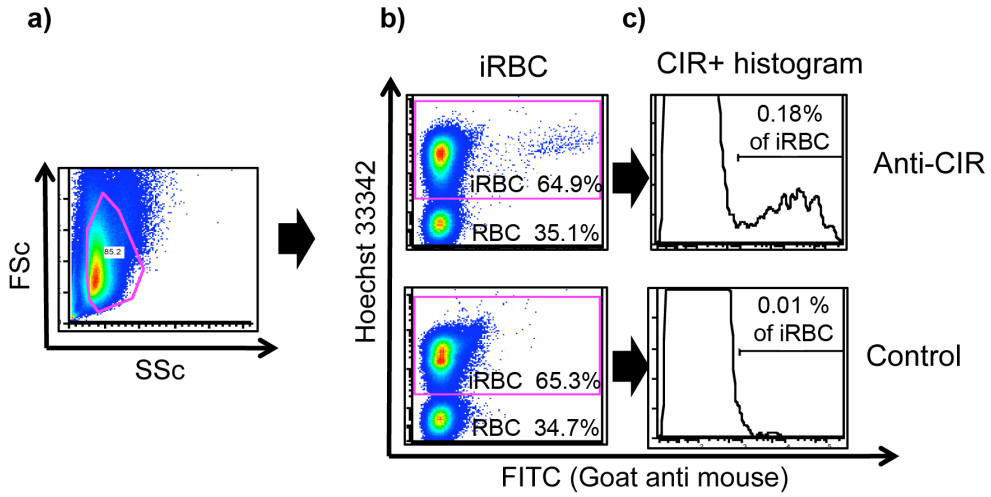


Figure 40: CIR surface localization using flow cytometry of live iRBC.

Live trophozoite stage *P. chabaudi* iRBCs were probed with: Hoechst 33342, for detection of parasite nuclei and mouse pre-immune serum (control) or anti-serum to the conserved CIR peptide (Anti-CIR), detected using a goat anti-mouse FITC conjugated secondary Ab.

Live iRBCs were defined first using forward scatter (FSc) and side scatter (SSc) measurements, shown within the pink region, **a)**. These cells were determined to be iRBCs, which had high levels of Hoechst 33342 staining, indicating the presence of nuclei, shown within the pink region, **b)**. Histograms show the proportion of iRBC which were CIR positive, determined by positive staining in the FITC channel and shown by the black horizontal bar, **c)**.

5.4 Discussion

The work described in this chapter was to detect CIR proteins expressed by *Plasmodium chabaudi* iRBCs, and their location within iRBCs. *P. chabaudi* proteins with a range of sizes were recognized by polyclonal anti-CIR Abs, which may represent several different CIR proteins expressed during this infection or the degradation products of one dominant CIR. All trophozoite stage iRBCs displayed CIR localization surrounding the parasite nuclei. At least half of trophozoite stage iRBCs also exhibited CIR staining close to the iRBC surface membrane. Flow cytometric analysis of live-stained iRBCs suggested that a small proportion of iRBCs may have expressed CIR at the surface. Alternatively, CIR proteins could be located just underneath the iRBC membrane, and therefore were not exposed to Abs during staining of live iRBCs.

VIR proteins have also been shown to localize near to the surface of trophozoite stage iRBC, via immuno-fluorescence analyses carried out using fixed *P. vivax* iRBCs (Del Portillo et al., 2001, Fernandez-Becerra et al., 2005). Exposure of YIR proteins at the iRBC surface has been confirmed using immuno-fluorescent staining of live, un-permeabilized iRBCs (Cunningham et al., 2005), which may also be true for other PIR family members. Janssen and colleagues have provided preliminary evidence for similar CIR localization, using Abs generated against a peptide with closest homology to the CIR PCHAS_000020 (Janssen et al., 2004). The work described in this chapter aimed to determine whether the localization observed for PCHAS_000020 was characteristic of the majority of CIRs, and was comparable to that observed for other PIR.

Immuno-fluorescence assays using fixed iRBCs, probed with Abs generated to the conserved CIR peptide, showed that CIR proteins were located around the parasite nuclei, probably within the PV. Confirmation of PV localization would require the use of a PV-specific marker, such as the monoclonal Ab MoAb 8E7/55, which recognizes *P. falciparum* EXP-1 (Lord et al., 1993), and an equivalent protein in *P. chabaudi* of 54 kDa (Sam-Yellowe, 2009). CIRs were also detected in discrete dots throughout the cytoplasm of trophozoite stage iRBCs. It is tempting to speculate that these dots could represent CIR proteins in the process of export to the iRBC surface. Vesicular structures such as the Maurer's clefts are thought to traffic proteins during export [reviewed by: (Sam-Yellowe, 2009)], and the Schüffner's dots of *P. vivax* also comprise caveolar vesicle complexes (Alkawa et al., 1975), which may form part of the protein export pathway. In fact, 28 kDa *P. vivax* antigens have been identified within the Schüffner's

dots (Matsumoto et al., 1988), which are within the predicted size range of the VIR proteins.

A small number of *P. chabaudi* iRBCs also appeared to have CIR proteins at the surface, detected using live-staining followed by flow cyometry. The small number of these detected iRBCs may be a consequence of the anti-CIR peptide Abs used. The conserved CIR peptide corresponds to motif 3 identified in chapter 2.3.3, which appears to be located within the hydrophobic protein core (chapter 4.2.4). As previously discussed, this peptide is likely to play a role in maintenance of CIR protein structure, and may therefore not be accessible to the anti- CIR peptide Abs when the proteins were not denatured (in SDS-PAGE, under reducing conditions) or permeabilized by fixation. This would explain the apparent discrepancy between the detection of CIR close to the iRBC membrane in the majority of iRBCs by fixed immuno-fluorescence, and the detection of less than 1% of iRBCs appearing to contain surface exposed CIR by live staining and flow cytometry.

Digestion of proteins exposed at the iRBC surface with trypsin could be used to verify the surface expression of CIR by mass spectrometry of the cleaved peptides. Alternatively, using a stronger digestion, such as proteinase K, all proteins would be lost from the iRBC surface. If the detection of a particular CIR following such treatment was then lost in comparison to un-treated control iRBCs, this would indicate that the CIR was indeed present at the iRBC surface.

In addition to confirming the localization of CIR proteins in trophozoite stage parasites, the data presented here provide novel evidence for the localization of CIRs in other iRBC stages. All merozoites contained CIR localization at the apical end, in contrast to the presence of MSP1 on the whole merozoite surface. This agrees with previous data finding *yir* transcription in merozoites (Shi et al., 2005). The pattern of CIR localization observed in merozoites is strongly reminiscent of that described for members of the RIFIN A sub-family (Narum and Thomas, 1994), which were located adjacent to AMA1 proteins, found at the apical end of the merozoite (Narum and Thomas, 1994).

This provides strong support to the similarities in CIR and RIFIN phylogenetic structure described in chapter 2.3.5. It is possible that the two major CIR sub-families identified in chapter 2 could also display differential cellular localizations, as observed for A and B type RIFINs (Narum and Thomas, 1994). A type RIFINs localize throughout the iRBC cytosol and at the iRBC surface in trophozoites, whilst B type RIFINs are found

only within the parasite. If anti-CIR Abs recognized both CIR sub-families, the detection of CIR within the PV and at the iRBC surface could have incorporated the differential localizations of both sub-families. To investigate whether this is the case, and whether *P. chabaudi* merozoites also display the different CIR localizations observed for RIFINs (Petter et al., 2007), reagents specific to each sub-family must be developed.

The idea of differential localization of PIR sub-family members has already been investigated for the VIR proteins to some extent. Del Portillo and colleagues generated Abs to peptides representing the sub-families A, C and E [KKKKRKKRNYYDYGWC, QKDWRERKALYDYC and RKFARNLKNISTILNDC, respectively (Fernandez-Becerra et al., 2005)]. Anti-C and E sera were used to investigate whether VIR proteins belonging to different sub-families displayed different cellular localizations. In schizonts, both VIRs appeared to be located in discrete points throughout the iRBC cytosol and each anti-serum produced an almost identical pattern (Fernandez-Becerra et al., 2005). The C- and E- specific Abs were shown not to recognize peptides belonging to the other families by Western blot (Fernandez-Becerra et al., 2005), nonetheless it is possible that they recognized other shared features of the different VIR sub-families. This illustrates an important caveat for the use of polyclonal anti-sera for such experiments; namely that a range of Ab specificities exists within the sera, some of which may recognize different sub-families.

To investigate the possibility that members of different PIR sub-families may perform different functions, reagents to detect PIR sub-family members must therefore be highly specific, recognising few variants. As described above, the use of polyclonal Abs for this purpose is limited due to the range of Ab specificities and avidities. In contrast, monoclonal Abs (mAbs) are mono specific, therefore the acquisition of mAbs to CIR proteins would be useful for wide-ranging applications, including confident protein localization, because any detected cross-reactivity would be due to the presence of shared epitopes within proteins, rather than their recognition by different Abs within the polyclonal pool. The specificity of monoclonal Abs may also be mapped to determine the exact epitope that is recognized, allowing determination of which CIRs are recognized, and providing information about CIR protein folding. In addition, as monoclonal Abs are generated from a cell line, different batches can be safely assumed to contain the same specificity, unlike poly-clonal sera, whose range of specificities is highly dependent upon the antigen batch used and / or different animals' responses.

An alternative approach for determination of CIR localization would be to genetically manipulate *P. chabaudi* by transfection with reporter constructs, such that CIR proteins within each major sub-family would be expressed along with a fluorescent tag. Fusion of BIRs to GFP has been used in two studies to investigate their cellular localization. Firstly, Di Girolamo and colleagues found that BIRs were associated with cholesterol rich membrane microdomains, exported into the iRBC cytosol and likely to the iRBC surface (Di Girolamo et al., 2008). Subsequently Sijwali and Rosenthal were able to use various BIR, RIFIN and STEVOR constructs, amongst others, to investigate the requirements for trafficking of these proteins (Sijwali and Rosenthal, 2010). This work also found that BIRs were exported to the iRBC cytosol and confirmed the localization of B-type RIFINs to the parasitophorous vacuole. It was suggested that the C-terminus of RIFIN and STEVOR proteins is important for their trafficking, since fusion proteins containing the *Pf*EMP1 N-terminus, GFP and the RIFIN C-terminus were exported to the iRBC cytosol (Sijwali and Rosenthal, 2010). The N-terminus of a STEVOR protein has also been shown to enable trafficking to the iRBC cytosol (Saridaki et al., 2009), so the crucial signals that enable iRBC localization of STEVOR and RIFIN proteins are yet to be confirmed.

Formerly, such transfection studies have not been feasible in *P. chabaudi*, however recent methodological advances have enabled the efficient generation of integrated transfectants expressing, for example, the reporter proteins M-cherry and luciferase (Spence et al., 2011). Following such improvement of *P. chabaudi* transfection methods, it should now be possible to tag CIR proteins in a similar way to the studies described above, allowing investigation of the cellular localization of individual CIR family members.

A note of caution however is that the fusion of a large fluorescent protein to CIR may alter the native protein trafficking [observed for example by (Sijwali and Rosenthal, 2010, Kuhn et al., 2010)]. An alternative to using a fluorochrome-CIR fusion protein would be to tag endogenous CIRs with a small tag, which could then be detected by IFA using Abs recognizing the tag. An example is the 10 amino acid TY1 tag, developed from the *Saccharomyces cerevisiae* Tyl virus-like particle (Bastin et al., 1996), which is specifically recognized by two different monoclonal Abs (Brookman et al., 1995) and has been used to confirm localization and interactions for several *P. falciparum* and *T. gondii* proteins (Haase et al., 2008, Michelin et al., 2009, Treeck et

al., 2009). Such approaches could be used to confirm the CIR localization observed here by IFA, following which, investigations into their possible function(s) may proceed.

A recent study by Carvalho and colleagues was able to make use of polyclonal Abs recognizing the A and E sub-groups (Fernandez-Becerra et al., 2005), to block the capacity for iRBCs to cyto-adhere to human lung endothelial cells by between 20 and 40% compared to the extent of cyto-adhesion of iRBCs observed in the absence of specific Abs (Carvalho et al., 2010). Their data suggest that VIR proteins may be involved in sequestration to some extent, although only two *P. vivax* isolates were investigated so the implication of VIR proteins in cyto-adhesion remains to be substantiated. If confirmed, a role in mediating cyto-adhesion may also be true for members of the CIR family, since *P. chabaudi* is known to cyto-adhere to CD36 and to sequester, particularly in the liver (Gilks et al., 1990, Mota et al., 2000).

To investigate the contribution of CIR proteins to the cyto-adhesion of *P. chabaudi* iRBCs, there are several possible approaches. Similar to the experiment described above in relation to VIR proteins (Carvalho et al., 2010), the anti-CIR Abs generated in this study could be used for inhibition of iRBC binding to mouse endothelial cells *in vitro*, using cell lines derived from known sites of sequestration, for example hepatic sinusoidal endothelial cells [which comprise the majority of endothelial cells in the liver, reviewed by (Huebert et al., 2010)], or to purified endothelial receptors such as CD36. Mota and colleagues have shown that Abs present in sera from *P. chabaudi* immune mice are able to disrupt the binding of iRBC both to CD36 and rat endothelial cells (Mota et al., 2000).

In addition, the recent technical advances in *P. chabaudi* transfection in this laboratory have enabled the generation of luminescent *P. chabaudi* parasites (Spence et al., 2011). Firstly, these parasites could be used directly to investigate the sequestration behaviour of iRBC *in vivo*, using whole body imaging of *P. chabaudi* infected mice. This technique has been applied to another rodent malaria parasite, *P. berghei*, to compare the sequestration phenotype between wild-type and CD36 deficient mice (Franke-Fayard et al., 2005). Use of luminescent *P. chabaudi* iRBCs to investigate the role of CIR proteins in sequestration could either use passive transfer of anti-CIR Abs to determine their effect on parasite sequestration, or the transgenic parasites could be further adapted to over-express particular CIR proteins either alone or in combination with luciferase. If this were possible, CIR over-expressing parasites could be used to initiate studies comparing the sequestration phenotype between parasites with high and

wild-type CIR expression both *in vitro* and *in vivo*. Over-expression of *cir* promoters may result in down-regulation of CIR expression, as has been observed for the *rif, stevor* and *var* genes (Howitt, Willnski et al. 2009); thus this strategy would require the use of strong, non-*cir* promoters to achieve high levels of CIR expression.

CIR protein expression could also be determined in other stages of parasite development, such as gametocytes, ookinetes, oocysts, sporozoites and infected hepatocytes. RIFIN and STEVOR expression have been detected within gametocytes and sporozoites (Florens et al., 2002, McRobert et al., 2004, Petter et al., 2008). BIR peptides have also been detected within gametocytes and sporozoites and additionally in oocysts and ookinetes (Hall et al., 2005). It is likely therefore, that CIR proteins would also be expressed in other stages of the lifecycle beyond asexual erythrocytic growth, which could be elucidated by using the anti-CIR Abs generated here for immunofluorescent studies of these parasite stages. The detection of CIR proteins in multiple stages of parasite development would indicate, first, that these proteins mediate effects required for parasite survival in different environments, which may include evasion of mosquito immunity by oocysts. Second, if different CIR family members were found to be associated with particular stages of parasite development and / or different cellular localizations, the 'sub-functionalization' detected within the CIR family (described in chapter 2.3.5) could mean that these proteins have evolved to fulfil multiple functions.

Chapter 6: Immune recognition of CIR proteins

6.1 Introduction

The detection of CIRs at the apical end of *P. chabaudi* merozoites and possibly at the surface of some iRBCs (Chapter 5), indicates that these proteins could be targets for Ab-mediated immunity during *P. chabaudi* AS infection.

Passive transfer of purified IgG from clinically immune adults has demonstrated that substantial protection against clinical malaria is afforded by Abs (Cohen et al., 1961, Edozien et al., 1962, McGregor and Carrington, 1963, Sabchareon et al., 1991). These results have been confirmed by experiments carried out in rodent models, including *P. chabaudi* (Diggs and Olser, 1969, Phillips and Jones, 1972, Boyle et al., 1982). In the model *P. chabaudi adami*, mice were unable to clear parasites in the absence of B cells (Grun and Weidanz, 1983, Meding and Langhorne, 1991, Rotman et al., 1998), indicating that Abs are crucial for resolution of infection and for immunity to re-infection. In addition, Abs appear able to mediate parasite clearance even in the absence of Fcγ dependent effector functions (Rotman et al., 1998), which may include agglutination and / or neutralization (as described in section 1.3.2). Antibody production also reflects the nature of CD4+ help available, as different IgG isotypes are generated within the particular cytokine environments evoked by Th1 and Th2 responses (Paul et al., 1987, Purkerson and Isakson, 1992, Collins and Dunnick, 1993, Else and Finkelman, 1998). Therefore the measurement of Abs and their isotype distribution reflects several aspects of the immune response against malaria, particularly in model systems [for example: (Fairlie-Clarke et al., 2010)].

The crucial antigen targets of protective Abs are still largely unknown however, and it is imperative to investigate a wide range of antigens for this reason. Abs involved in parasite clearance are thought to target antigens expressed on the surface of merozoites or iRBCs, including ligands involved in iRBC invasion (Miller et al., 1975, Brown, 1976, Cohen, 1979, Quinn and Wyler, 1979, Freeman et al., 1980, Jarra and Brown, 1989, Snounou et al., 1989, Mota et al., 2001, Bull and Marsh, 2002, Good et al., 2004). The majority of studies investigating immune responses to *Plasmodium* have focussed on the recognition of conserved proteins, for example the merozoite stage antigens MSP1 and AMA1 which are not antigenically variant within a parasite strain. Both AMA1 and MSP1 are known to be immune targets (Egan et al., 1999, Hodder et al.,

2001, O'Donnell et al., 2001), and several studies have shown a correlation between high titres of anti-AMA1 and / or MSP1 antibodies and protection from clinical malaria (Weir and Cockerham, 1984, Riley et al., 1992, Egan et al., 1996, Su and Wellems, 1996, Branch et al., 1998, Anderson et al., 1999, Dodoo et al., 1999, Kitua et al., 1999, Conway et al., 2000, Metzger et al., 2003, Polley et al., 2004, Perrant et al., 2005, Nebie et al., 2008, Stanisic et al., 2009, Osier et al., 2010).

Passive transfer of monoclonal anti-MSP1 and anti AMA1 Abs has been able to substantially reduce parasitaemia in rodent models (Boyle et al., 1982, Gozalo et al., 1998, Narum et al., 2000, Vukovic et al., 2000, Lozano et al., 2011), and immunization studies using these antigens have also been able to generate protective immunity (Daly and Long, 1993, Ling et al., 1994, Crewther et al., 1996, Hirunpetcharat et al., 1997, Tian et al., 1997, Anders et al., 1998, Burns et al., 2003, Burns et al., 2004, Alaro et al., 2010). For these reasons both antigens have been developed as vaccine candidates, although substantial obstacles to their efficacy have emerged. Not only are high titres of specific Abs required for protection (Hirunpetcharat et al., 1997, Ling et al., 1997), but the immunogenicity and activity of induced Abs can be substantially affected by the vaccine design [for example: (Hussain Reed et al., 2009)], delivery mechanism and adjuvant used [reviewed by: (Coler et al., 2009, Dey and Srivastava, 2011)]. In addition the correlation between naturally occurring Abs to AMA1 and MSP1 and protection is not completely clear (Dodoo et al., 1999, Sakihama et al., 1999, Corran et al., 2004, Okech et al., 2004), suggesting that immune responses, including the generation of Abs, to other antigens are also critical for development of protective immunity.

Less is known about Ab responses to VSAs, although these are also thought to be targets of protective immunity (Marsh et al., 1989, Bull et al., 1998). The acquisition of Abs against *Pf*EMP1 variants correlates with protection from severe malaria in Africa (Dodoo et al., 2001, Ofori et al., 2002, Kinyanjui et al., 2004b, Yone et al., 2005, Magistrado et al., 2007). In particular, the acquisition of Abs recognizing sub-type A *Pf*EMP1 variants appear particularly associated with protection against severe disease (Jensen et al., 2004, Bull et al., 2005b, Kaestli et al., 2006, Cham et al., 2009, Cham et al., 2010), whilst the presence of Abs to conserved regions of VAR2CSA correlate with immunity to PAM (Elliott et al., 2007, Rogerson et al., 2007). By contrast, there is no clear association of anti-RIFIN or STEVOR Abs with protection; with either no correlation to protection or perhaps even a link to developing a higher burden of parasitaemia or cerebral malaria (Abdel-Latif et al., 2003, Schreiber et al., 2006,

Schreiber et al., 2008). It is possible that the acquisition of Abs against the vast number of VSAs simply reflects the number of infections experienced, leaving much scope for understanding the connection between acquisition of Abs towards different VSA families and protection, either from severe disease or symptomatic infection.

Whilst indirect evidence indicates that PIR expression is modulated by the presence of an intact immune response (Cunningham et al., 2005), only two studies to date have attempted to investigate immune recognition of PIR proteins. Both studies were carried out in northern Brazil using sera from a cohort of *P. vivax* symptomatically infected immigrant workers (Fernandez-Becerra et al., 2005). The first study used 22 recombinant VIR fusion proteins corresponding to the second *vir* exon, and found differential recognition of these by immune sera from 32 patients. Approximately three quarters of patient serum samples were able to recognize at least one recombinant VIR by Western blot (Fernandez-Becerra et al., 2005), showing that VIR proteins are immunogenic during natural infection. The second study used seven of the recombinant VIR proteins and found VIR recognition by the immune sera of half of the 200 patients tested (Oliveira et al., 2006), the disparity was probably due to the lower number of recombinant VIR fragments tested in the second study (Oliveira et al., 2006).

In both studies, the sera of patients with a primary *P. vivax* infection contained Abs recognizing a range of VIR proteins, including those from different phylogenetic sub-families. There was no significant difference in the recognition of any VIR proteins by the sera of patients infected for the first time and those who had experienced several infections (Oliveira et al., 2006). Likely explanations for this phenomenon are that many VIR proteins are expressed during a primary *P. vivax* infection, and that Abs are able to recognize many VIRs. Transcriptional analyses of the *pir* family support the former scenario, whilst support for both hypotheses comes from the first study, where substantial cross-reactivity was observed between different VIR sub-families when polyclonal antisera were raised to different VIR peptides (Fernandez-Becerra et al., 2005). The expression of multiple VIRs during infection and the induction of cross-reactive Abs by VIR proteins may both contribute to the observed range of anti-VIR Abs induced in response to a primary *P. vivax* infection.

The limited access to patient material and stochastic nature of natural infection means that the contribution of immune responses against PIR proteins is difficult to analyze in the field. Rodent models such as *P. chabaudi* AS allow the detailed investigation of immune recognition of PIR proteins, under controlled conditions including the ability to

initiate *Plasmodium* infections with a defined parasite population. Objectives of the present study were to:

i) Confirm whether CIR proteins were recognized by immune sera, as previously observed for VIR (Fernandez-Becerra et al., 2005).

ii) Investigate whether the presence of high anti-CIR Ab titres, induced by immunization, would influence the outcome and the parasite behaviour during a subsequent *P. chabaudi* infection.

6.2 Methods

Origins and housing of parasites and mice have been described previously, along with measurement of parasitaemia during infection, chapter 3.2.4.

6.2.1 Preparation of *P. chabaudi* immune-serum

Serum samples were collected from 10 mice experiencing one, two or four *P. chabaudi* AS infections, normal (control) serum was collected from uninfected mice. Primary *P. chabaudi* AS infections were initiated intra-peritoneally, using 10^5 iRBC from a frozen stabilate that was expanded in $RAG2^{-/-}$ mice. Mice experiencing more than 1 infection were re-infected with 10 fold more parasites at each infection. Blood was allowed to clot at room temperature for 2 hours. Low speed centrifugation (1500 x *g* for 10 minutes) was then carried out and the clot removed, and the serum subjected to high-speed centrifugation (13,000 x *g* for 5 minutes) to remove aggregated material and free erythrocytes.

6.2.2 SDS PAGE and immuno-blotting

Proteins were resolved under reducing conditions using NuPAGE gels (Invitrogen), as previously described, chapter 4.2.3. 200 ng of the recombinant CIR proteins PCHAS_000100, PCHAS_070130 and PCHAS_040110 and $MSP1_{p21}$ were loaded, as determined by BCA assay, described in chapter 4.

All gels were used for immuno-blotting, where the proteins were electrophoretically transferred to Hybond C membrane (Amersham Biosciences) in buffer containing 10% methanol (v/v), 0.025 M Tris base (Sigma) and 0.2 M Glycine (BDH biosciences) for 3 h at 30 V. Membranes were blocked at 4 ˚C overnight in PBS containing 0.1% Tween and 3% BSA, and washed for 10 minutes three times with PBS containing 0.01% Tween 20.

The recombinant proteins were then probed with serum (diluted 1/50) from a naïve mouse or from a mouse which had experienced 1, 2 or 4 *P. chabaudi* AS infections. Antibodies directed against the histidine tag (Novagen), present on all recombinant proteins were used at 0.5 μg/ml concentration, as described in chapter 4.2.3 to ensure equal amounts of the recombinant proteins had been loaded.

Bound Abs were detected using horseradish peroxidase (HRP)-conjugated goat anti-mouse IgG, (Southern Biotechnology, used at 1:10,000 dilution) as the secondary Ab. Membranes were washed as before, and SuperSignal West-Pico chemi-luminescent substrate (Thermo Scientific) added, according to manufacturer's instructions, and incubated in the dark for 5 minutes. Light emission was detected by exposure of the blot to Biomax film (Kodak).

6.2.3 Enzyme linked immuno-sorbent assays (ELISAs)

The western blots described above surveyed a very limited number of CIR proteins, resolved under reducing conditions, which may have disrupted conformational determinants recognized by anti-CIR Abs present in the sera of *P. chabaudi* immune mice. Thus, ELISAs were carried out using these proteins, and with peptides representing a larger number of CIRs. The methodology used for these assays was the same as described in chapter 5.2.4, with the following exceptions:

To measure CIR-specific Abs present in *P. chabaudi* immune sera, 96-well flat Polysorb plates (Nunc, Denmark) were coated with 50 µl of 5 µg/ml recombinant MSP-1, CIR proteins or BSA conjugated CIR peptides (for which the approximate total protein coat for the conserved CIR peptide, CIR sub-family U peptide and CIR sub-family A-F peptide respectively was 3.75 µg, 1.09 µg and 0.93 µg per well when the BSA conjugate was accounted for). The conserved CIR peptide was insoluble alone hence the BSA-conjugate was required for binding to the ELISA plate. For comparability, the other peptides were treated in the same manner. Antigens were diluted in PBS, and 50 µl was applied per well. Plates were incubated overnight at 4 ˚C, following which, the residual surface was blocked with 200 µl of blocking buffer (PBS containing 1% BSA, 0.3% Tween 20 and 0.05% Sodium azide). Recombinant $MSP1_{21}$, CIR proteins and peptides were probed with sera diluted 1/50 in blocking buffer from ten naïve mice or from ten individuals which had experienced 1, 2 or 4 *P. chabaudi* AS infections. The titres of specific Abs induced by immunization were calculated by interpolation of the titration curve, using non-linear regression analysis performed in Prism v5.0. The titre of Abs was defined as the dilution at which the sera from each *P. chabaudi* infected mouse would reach the OD405 nm of *P. chabaudi*-naïve mouse sera.

To measure a known quantity of each antigen-specific Ig isotype, 96-well flat Polysorb plates (Nunc, Denmark) were coated with 50 µl of 5 µg/ml CIR sub-family U or A-F

specific peptides (not conjugated to a carrier protein, as these peptides were soluble in PBS). Each plate was also coated with 50 µl of a 5 µg/ml capture goat anti-mouse Ig (Southern Biotechnology), which may bind purified Ig of each isotype, used as a standard (Langhorne et al., 1989, Quin and Langhorne, 2001, Hensmann et al., 2004). Antigens were diluted in PBS, and 50 µl was applied per well. Plates were incubated overnight at 4 ˚C, following which, the residual surface was blocked with 200 µl PBS containing 1% BSA, 0.3% Tween 20 and 0.05% Sodium azide. CIR peptides were probed with sera diluted 1/50 in blocking buffer from a pool of ten naïve mice or from ten individuals which had experienced 4 *P. chabaudi* AS infections. In addition, a range of dilutions (in blocking buffer) of purified mouse IgM, IgG1, IgG2a, IgG2b, and IgG3 (Sigma-Aldrich) were used to probe the capture Ab, as a standard curve. These bound antibodies, along with isotypes of Abs binding to CIR proteins on the plate were then detected using alkaline phosphatase-conjugated goat anti-mouse IgM, IgG1, IgG2a, IgG2b or IgG3 Abs (Southern Biotechnology). Antigen-specific Abs were ennumerated relative to the standard by interpolation of the titration curve, using non-linear regression analysis performed in Prism v5.0.

In addition, ELISAs were also carried out to determine the titres of antigen-specific Abs induced by immunization of mice with recombinant $MSP1_{21}$, CIR proteins and peptides, described in section 6.2.4.

6.2.4 CIR-immunization and *P. chabaudi* challenge

Animals were immunised with 50 µg of antigen in the presence of 12 µg Abisco adjuvant (Iscanova). Antigens used were the recombinant proteins $MSP1_{p21}$, CIRs PCHAS_040110 and PCHAS_000100 (generated in chapter 4); the conserved CIR peptide (described in chapter 5, Figure 36), and CIR sub-family U and A-F peptides (described in chapter 6.3.1, Figure 42). Here it must be noted that immunization with the CIR peptides required a carrier protein. For this purpose all peptides were conjugated to the carrier proteins BSA or KLH by Jerini Peptide Technologies, Berlin, Germany. Unfortunately, turbidity of the KLH conjugated peptide solutions prevented estimation of an accurate concentration. Instead, the BSA conjugated peptides were used for immunization as an accurate peptide dose could be administered.

Mice immunized with CIR peptides received 50 µg of peptide, as peptides were conjugated to the carrier protein BSA, the total protein dose in these animals was

greater than 50 µg (approximately 747.96 µg, 217.23 µg and 185.50 µg, for mice immunized with the conserved CIR peptide, CIR sub-family U peptide and CIR sub-family A-F peptide respectively). In addition, a group of mice were immunized with a CIR cocktail, who received 10 µg of each CIR antigen (CIR PCHAS_040110, PCHAS_000100, the conserved CIR peptide, and both CIR sub-family specific peptides) thus receiving 50 µg of CIR antigens in total (which equates to approximately 330 µg of total protein). Immunizations were administered three times at 2-week intervals into different sites: intra-muscular (left hind leg), intra-muscular (right hind leg) and sub-cutaneous.

The titres of antigen-specific Abs induced by immunization of mice with recombinant $MSP1_{21}$, CIR proteins and peptides were determined by ELISA. ELISAs were carried out according to the methods described above (in sections 5.2.4 and 6.2.3), with certain exceptions, which are described below.

96-well flat Polysorb plates (Nunc, Denmark) were coated with 50 µl of 5 µg/ml recombinant MSP-1, CIR proteins or BSA conjugated CIR peptides (for which the approximate total protein coat for the conserved CIR peptide, CIR sub-family U peptide and CIR sub-family A-F peptide respectively was 3.75 µg, 1.09 µg and 0.93 µg per well when the BSA conjugate was accounted for). Antigens were diluted in PBS, and 50 µl was applied per well. Plates were incubated overnight at 4 ˚C, following which, the residual surface was blocked with 200 µl blocking buffer, as described above. Recombinant $MSP1_{p21}$, CIR proteins and peptides were probed with serum samples from each mouse: compared before and after immunization. Naïve sera were diluted 1/1000, whilst immunised sera were diluted 1/1000, 1/10000 and 1/100000, in blocking buffer. CIR-specific Abs were detected using alkaline phosphatase-conjugated goat anti-mouse Ig (Southern Biotechnology). The titres of Abs induced by immunization were calculated as described above, by interpolation of the titration curve, and were defined as the dilution at which the sera from each immunized mouse would reach the OD405 nm of sera from the same individual prior to immunization.

Detection of peptide specific Ab responses would ideally have been carried out by ELISA using unconjugated peptides as the coating antigen, however the conserved CIR peptide was insoluble in PBS. Instead, the BSA-conjugated peptides were used as coating antigens for ELISA as described below. This is shown in Figure 41a.

In order to determine the relative titres of Abs recognizing the peptides, rather than the carrier protein BSA in the sera of CIR peptide immunized mice, purified BSA (Sigma Aldrich), or BSA conjugated CIR peptides were used as coating antigens for ELISA. Solutions containing the same molarity of BSA were prepared in PBS and used to coat the wells, so the antigens would be approximately comparable. The molarity used was the equivalent of a 5 μg/ml BSA solution, 74.88 nM, and 50 μl was applied per well. ELISAs were otherwise performed as described above, for determination of the titres of CIR-specific Abs induced by immunization, Figure 41b.

Finally, to confirm that immunization with the CIR peptides did result in the generation of Abs recognizing the peptides, rather than the carrier protein BSA, the PBS soluble unconjugated CIR sub-family specific peptides were used as coating antigens for ELISA. The antigens were prepared at 5 μg/ml concentration in PBS and 50 μl applied per well. ELISAs were otherwise performed as described above, for determination of the titres of CIR-specific Abs induced by immunization, shown in Figure 41c.

Whilst these three methodologies were different and thus not able to provide an exact calculation, three points may be observed. First, the sera of all groups of mice immunized with BSA-conjugated CIR peptides had similarly high titres of Abs recognizing their cognate antigen, crossing the background OD405 nm of pre-immune sera at titres of approximately $1x10^5$, Figure 41a. Second, the sera of all mice had higher titres of Abs recognizing the BSA-conjugated peptide than to BSA alone, Figure 41b. From this it may be inferred that the majority of Abs were generated against the CIR pepides. Finally, for the two peptides that were soluble in PBS, it was possible to confirm that titres of Abs recognizing the un-conjugated peptide were of a similar titre to those generated against the BSA conjugated peptide, of $1.28\ x10^5$, Figure 41c. This was likely to also be the case for the mice immunized with the PBS insoluble, conserved CIR peptide.

One important caveat here was that BSA was used both as a carrier protein for immunization and to block the ELISA plates. Whilst Abs were diluted in an excess of BSA to remove BSA-reactive Abs, it is possible that some BSA-reactive Abs were able to bind to the BSA block. ELISAs were thus carried out using the same strategy for the sera of all mice immunized with CIR peptides, with un-conjugated CIR sub-family specific peptides as coating antigens (Appendix 6.3).

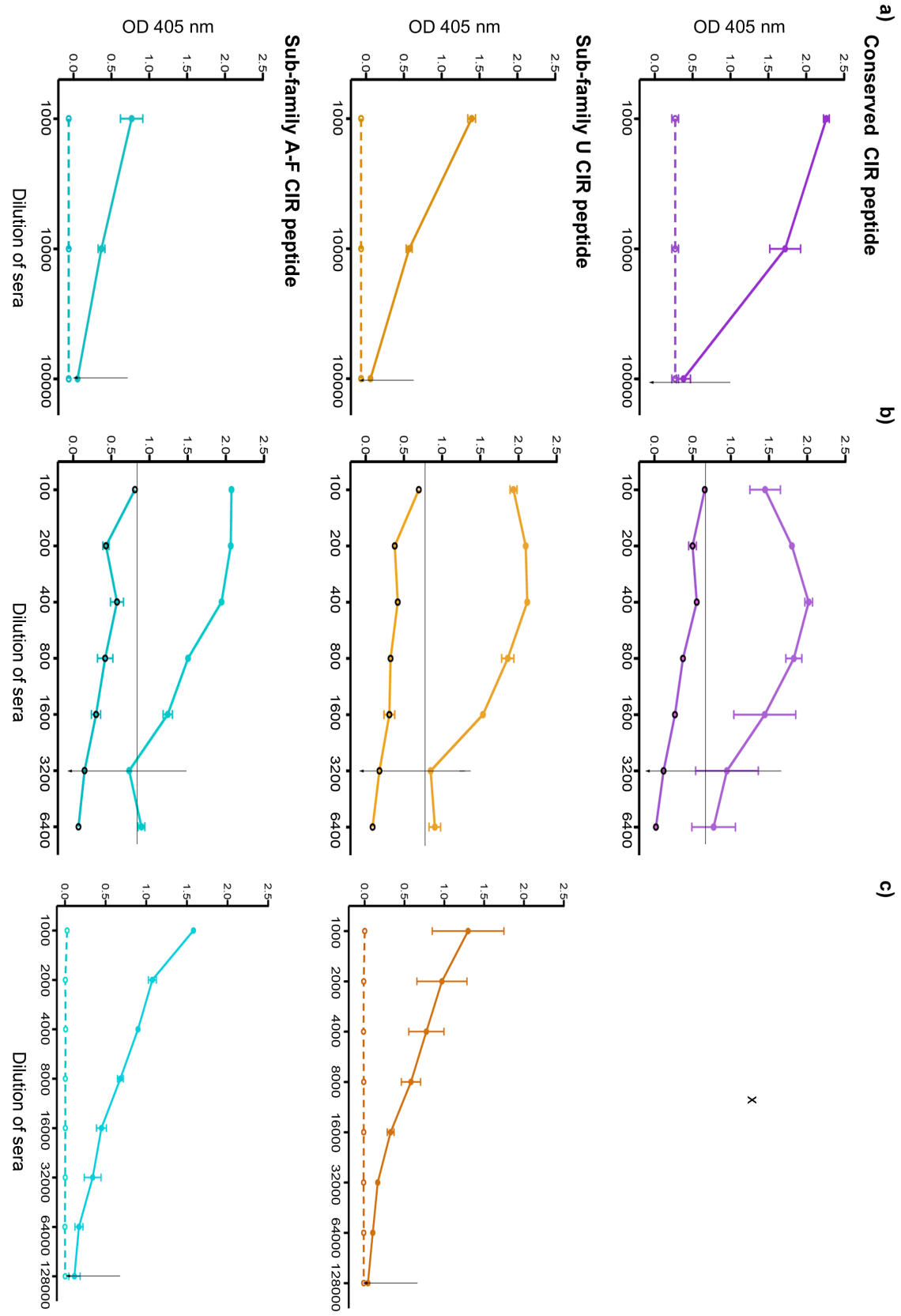

a)
b)
c)
Conserved CIR peptide
Sub-family U CIR peptide
Sub-family A-F CIR peptide
OD 405 nm
Dilution of sera
x

Figure 41: Titres of Abs induced after immunization with CIR peptides, conjugated to the carrier protein BSA.

Sera from mice immunized with the conserved CIR peptide (purple), CIR sub-family U peptide (orange), or CIR sub-family A-F peptide (blue), were compared with pre-immune sera from the same animals.

The mean absorbance OD405 nm of the ELISA substrate is plotted against the dilution of sera, from triplicate titrations. Bars represent the standard error of the mean. The sera of mice immunized with CIR peptides are represented by filled symbols and solid lines, whilst pre-immune sera are represented by empty symbols and dashed lines.

The coating antigens used for ELISAs were: BSA-conjugated CIR peptides, **a)**, BSA-conjugated CIR peptides and BSA alone, **b)**, or unconjugated CIR peptides, **c)**.

The only Abs recognizing each CIR peptide were found in the sera of mice immunized with those same peptides (Appendix 6.3). Thus, it was concluded that any BSA reactive Abs had been absorbed on the excess BSA during dilution of the sera. Therefore, although Abs were also generated against the carrier protein BSA, mice did make Abs recognizing the CIR peptide with which they were immunized, at least for the CIR sub-family specific peptides.

Control mouse groups of the same age were immunised with either saline alone, or saline plus adjuvant, using the same routes of administration. Before immunisation, pre-immune serum was collected from all of these animals. Twelve days after the last immunisation, sera were collected from clotted blood, aliquotted and stored at -20 ˚C.

Two days later, all mice were infected intra-venously with 10^3 *P. chabaudi* AS from a frozen stabilate that was expanded in a $RAG2^{-/-}$ mouse. Infections were monitored by daily blood smear. At the peak of infection (day 8), 50 µl blood was taken from each mouse, collected in 100 µl Krebs saline, containing 25 U/ml heparin (Leo Laboratories). To each sample 250 µl Trizol (Invitrogen) was added, and material stored at -80 ˚C.

This material was processed for RNA using phenol-chloroform extraction followed by cDNA synthesis and RT-qPCR, carried out as described in chapter 3.2 (Appendix 3.4), using *cir* primers listed in Table 12. Transcription of *cir* genes was analyzed for three mice from each group immunized in the presence of Abisco adjuvant. Ratios of *cir* transcription were calculated relative to the donor $RAG2^{-/-}$ mouse and normalized to *beta tubulin*. These data are tabulated in Appendix 6.7.

6.3 Results

6.3.1 Reagents used for detection of anti-CIR responses

In order to detect anti-CIR responses that may occur during *P. chabaudi* infection, CIR protein reagents were required. The recombinant CIR proteins PCHAS_000100 and PCHAS_040110, generated in chapter 4, were expressed in sufficiently high quantities that they could be used for this purpose. These proteins represented each of the major sub-families identified in chapter 2: U (PCHAS_040110) and A-F (PCHAS_000100). Although neither *cir* gene was dominant in the samples analyzed by RNA sequencing (Figure 23), each *P. chabaudi* infection appeared to have a different dominant *cir* transcript, and RT-qPCR analyses had shown that both genes were expressed during *P. chabaudi* infection (Figure 21 and Figure 22).

In addition, the identification of conserved and sub-family specific amino acid motifs within the CIR repertoire in chapter 2 provided a route for investigation of CIR proteins, by the synthesis of CIR peptides. Three amino acid motifs were chosen for peptide design. The first of these has been described in chapter 5, representing a highly conserved amino acid sequence expected to represent the majority of CIRs, illustrated in Figure 36. In addition, a motif specific to each of the major CIR sub-families was chosen for peptide synthesis, to enable functional comparisons between the CIRs and whether there were further similarities with the RIFINs of *P. falciparum*. The peptides were designed using the CIR sub-family U amino acid motif (identified in Figure 20d), and a motif found only in CIR sub-families A-F (motif 11, Table 7). These peptides are depicted in Figure 42.

6.3.2 CIRs are recognized by immune sera

To investigate the Ab response to CIR proteins following *P. chabaudi* AS infection, sera were prepared from mice experiencing one or more *P. chabaudi* infections. CIR-reactive Abs were detected using recombinant proteins and peptides representing different CIR sub-families and regions of the proteins, as described above. The anti-CIR Ab response was compared with the Ab response to the C-terminal fragment of MSP1, a protein coded for by a single copy gene, and known to induce Ab responses in a *P. chabaudi* infection [for example (Langhorne et al., 1989, Quin and Langhorne, 2001,

Hensmann et al., 2004).]. Recombinant $MSP1_{p21}$ was generated within the same expression system as the recombinant CIRs, using the same C-terminal His-tag (Hensmann et al., 2004).

P. chabaudi infection in mice induced Abs which recognized more than one of the recombinant CIRs as determined by Western blot, even after a single infection, Figure 43. After two infections Abs were present that recognized all CIR proteins, albeit at lower levels than anti- $MSP1_{p21}$ Abs. Recognition of all proteins increased after four infections.

The titres of CIR specific Abs were estimated from the sera of 10 mice using ELISAs, Figure 44. In addition to the recombinant proteins, three CIR peptides were used as coating antigens (described in chapter 2.3.6). The conserved CIR peptide represents 64% of CIR proteins, as revealed by BLAST analysis, whilst CIR sub-family U and A-F peptides were exclusively found in each of the major sub-families identified in chapter 2.3.2. Titres of Abs recognizing $MSP1_{p21}$, recombinant CIR PCHAS_040110 and the CIR sub-family U and A-F specific peptides significantly increased after four *P. chabaudi* infections, compared to the sera of mice experiencing only one infection, Figure 44.

In addition, Figure 44 shows that the titre of Abs recognizing any CIR proteins or peptides were approximately 100-fold lower after each infection than those from the same mice recognizing $MSP1_{p21}$. These differences were significant at the 99.9% confidence level, when assessed by one-tailed Mann-Whitney tests. The lower quantities and greater range of anti-CIR Abs observed in comparison to anti- $MSP1_{p21}$ Abs were likely to be due to the differential expression of CIR variants within each mouse's *P. chabaudi* infection. This is unlike MSP1, which is expressed from a single gene, present in all iRBCs.

Since antibody responses to the sub-family specific CIR peptides were reproducibly higher than those to both recombinant CIR proteins, these peptides were used for investigation of the isotype distribution of anti-CIR Abs in mice that had experienced four *P. chabaudi* AS infections. Figure 45 shows that the Ab isotypes recognizing these peptides were predominantly IgM & IgG2a, with no differences observed in the distribution of isotypes recognizing either CIR peptide.

6.3.3 CIR-immunization and *P. chabaudi* challenge

To investigate whether pre-existing anti-CIR Abs at high concentration might affect the course of a subsequent *P. chabaudi* AS infection (also referred to as challenge), an immunization experiment was carried out. The experimental design is shown in Figure 46. Groups of five mice were immunized three times with 50 μg of CIR proteins or peptides, $MSP1_{p21}$, or saline, in the presence of 12 μg Abisco adjuvant (Iscanova). In addition a negative control of saline without adjuvant was performed to enable detection of adjuvant-based effects. Fourteen days after the final immunization, mice were infected with *P. chabaudi* AS. Parasitaemia was monitored by thin-blood smear every other day after the first observation of parasites, and also at the peak of parasitaemia.

Several key variables were measured, outlined below:

- Anti-CIR Ab titres after immunization, to ensure that responses were made to the appropriate antigen in each individual.
- Parasitaemia during *P. chabaudi* challenge, to measure the potential effect of CIR-immunization upon parasite growth.
- Patterns of *cir* transcription at the peak of parasitaemia, to investigate likely changes in the population of parasites and their transcriptional responses in the presence of anti-CIR Abs.

Before carrying out the immunization and challenge experiment, it was important to to confirm that the CIR proteins and peptides were capable of inducing Abs that recognized native *P. chabaudi* proteins. As described in chapter 5 for the conserved CIR peptide, rabbits were immunized with each CIR protein or peptide for the generation of polyclonal anti-sera. These sera were used to confirm that immunization with any of the recombinant CIR protein or CIR peptides contained Abs capable of recognizing native *P. chabaudi* proteins by Western blot (Appendix 6.1).

6.3.3 i) CIR immunization and P. chabaudi *challenge*

Following immunization of mice for the challenge experiment, serum from each mouse was compared to pre-immune sera from the same individuals by ELISA. Both saline control groups had no Abs recognizing antigens above a threshold of approximately 1 x 10^3, thus sera containing titres above this level were deemed to contain Abs recognizing CIR proteins or peptides, or $MSP1_{p21}$ (Appendix 6.4). The titres of Abs induced by CIR

immunization are shown in Figure 47. The standard errors of the mean are shown, but were masked by the log scale of the graphs.

Sera from mice immunized with PCHAS_040110 contained similar titres of Abs recognizing both recombinant CIR proteins (titres of approximately 1 x 10^6), shown in the left panel of Figure 47a. Some cross-reactivity could potentially have been due to cross-reactive recognition of the His tag present at the C-terminus of both proteins; however, there was no recognition of the same His tag present on recombinant $MSP1_{p21}$. Instead it is likely that the two recombinant CIRs shared several conformational determinants, which could be recognized by Abs induced by immunization with either protein. Upon *P. chabaudi* challenge, mice immunized with PCHAS_040110 displayed a significant difference from saline-controls on day 7 of infection, assessed by one-tailed Mann Whitney test, shown in the right panel of Figure 47a.

The sera of mice immunized with the CIR sub-family U peptide had high titres (of approximately 1 x 10^7) of Abs to this peptide, Figure 47b. Surprisingly, Abs were also detected that recognized the conserved CIR and sub-family A-F peptides, although at low titres (less than 1.6 x 10^4). As immunization with all CIR peptides used BSA conjugates, there was a strong likelihood that such apparent cross-reactivity was in fact due to Ab recognition of the carrier protein by ELISA. To confirm that this was the case, ELISAs were carried out to measure the recognition of BSA alone and, where possible, to determine anti-peptide Ab titres using un-conjugated peptides. This was addressed in detail in section 6.2.4, where specific responses were detected in all cases to the cognate peptide (Figure 41). In addition, ELISAs were carried out to measure the degree to which sera from mice immunized with each CIR peptide recognized either of the un-conjugated peptides (Appendix 6.3). No cross-reactivity was detected; hence all apparent cross reactivity between the sera of mice immunized with CIR peptides (presenting as Ab titres less than or equal to 1 x 10^4) could be attributed to recognition of the carrier protein BSA.

Recognition of PCHAS_040110 by the sera of CIR sub-family U peptide immunized mice however, was likely not to be an artefact but recognition of the exact sub-family U peptide sequence: LPELTLPEGLYDCKTE, found within this recombinant CIR protein. Despite high titres of Abs recognizing the CIR sub-family U peptide and some cross-reactivity with PCHAS_040110 in these mice, there was no difference in parasitaemia compared to the control groups during *P. chabaudi* challenge infection, Figure 47b.

Similarly to mice immunized with the recombinant CIR PCHAS_040110, mice immunized with recombinant PCHAS_000100 also made Abs recognizing the other recombinant CIR. However, the mice immunized with PCHAS_000100 made 100 fold higher titres of Abs to the protein they were immunized with than to PCHAS_040110 (1 x 10^7 and 8.5 x 10^4 respectively), Figure 47c. In addition, low-level responses were detected against the sub-family A-F peptide (less than 1x 10^4), which is likely due to recognition of the peptide by Abs generated to a similar sequence within PCHAS_000100: KLSVDFYDSYIERNEYY (identical residues and those with similar properties are highlighted). Upon *P. chabaudi* challenge, mice immunized with PCHAS_000100 displayed the most significant differences in parasitaemia of all the groups tested, compared to saline- immunized controls. On day 3 and 7 of infection PCHAS_000100 immunized mice had significantly higher parasitaemia, followed by significantly reduced parasitaemia on days 11 and 13 of infection, Figure 47c. This suggests that the immunized mice were initially less able to control the growth of parasites, but after the peak of parasitaemia the mice were able to clear parsites slightly faster than control mice.

Sera from mice immunized with the CIR sub-family A-F peptide contained high titres of Abs to this peptide (1 x 10^6), shown in Figure 47d. A low level response was detected against the sub-family U peptide, Figure 47d, which was actually directed against the carrier protein BSA as described above (Figure 41, Appendix 6.3). In addition, mice immunized with the CIR sub-family A-F peptide made low titres of Abs (less than 1.4 x 10^4) recognizing both recombinant CIR proteins. As described above, PCHAS_000100 contained a sequence similar to this peptide, however the reason for recognition of the recombinant CIR PCHAS_040110 by mice immunized with a peptide not present in this protein is unclear. No differences were detected during the course of parasitaemia upon *P. chabaudi* challenge, Figure 47d, although one mouse died after the peak of infection, suggesting that these mice may have suffered from more acute pathology.

Mice immunized with the conserved CIR peptide were expected to generate Abs recognizing the majority of CIR proteins. The sera of these mice contained high titres (of 1 x 10^6) of Abs to this peptide, Figure 47e; and low-level responses to the sub-family A-F peptide (less than 1x 10^4), which were actually directed against BSA (Figure 41, Appendix 6.3). Similarly to the mice immunized with the CIR-sub-family A-F

specific peptide, no differences were detected during the course of parasitaemia upon *P. chabaudi* challenge, yet one mouse died after the peak of infection, Figure 47e.

Finally, the sera of mice immunized with the CIR-cocktail made low titres (less than 3.5 x 10^4) of Abs to each CIR antigen, Figure 47f. These mice displayed a significantly higher parasitaemia upon day 7 of *P. chabaudi* challenge than saline controls, (Figure 47f) similarly to the mice immunized with PCHAS_040110. Since PCHAS_040110 was present in the CIR cocktail, both groups of mice may have experienced Ab: parasite interactions leading to increased parasitaemia just prior to the peak of infection.

For comparison, a group of mice was immunized with recombinant $MSP1_{p21}$, as indicated in Figure 46. Like the recombinant CIR proteins, $MSP1_{p21}$ was generated in the same expression system, *P. pastoris*, and contained an identical C-terminal hexa-histidine tag for protein purification and identification (Hensmann et al., 2004). Mice immunized with recombinant $MSP1_{p21}$ made high titres of Abs (1 x 10^8) only recognizing this protein, yet these mice experienced no difference in parasitaemia upon *P. chabaudi* challenge, compared to saline controls (Appendix 6.4).

Due to the scope of this immunization and challenge experiment, measurement of pathology during the *P. chabaudi* challenge infection was impractical. A comparison of anaemia was however undertaken at the beginning and at the peak of infection, to determine whether any of the groups appeared more susceptible to this complication of malaria. No difference in anaemia was detected between groups (Appendix 6.5), except for the mice immunized with PCHAS_000100, whose rapid increase in parasitaemia after infection resulted in anaemia prior to the other groups.

6.3.3 ii) Measurement of cir *transcripts by RT-qPCR*

P. chabaudi infected blood samples were taken for analysis of *cir* transcription by RT-qPCR both from the donor $RAG2^{-/-}$ mouse, and from all challenged mice at the expected peak of parasitaemia (day 8 post infection). The ratio of *cir* transcript levels relative to the donor mouse was calculated for three mice per group and normalized to *beta tubulin* as described in chapter 3. The three mice chosen for this analysis were those whose parasitaemias were closest to the median parasitemia of the group on day 8 of *P. chabaudi* challenge.

The majority of *P. chabaudi* challenged mice displayed no difference in *cir* transcript levels compared to the donor mouse, with ratios of *cir* transcription being close to 1, shown in Figure 48. All RT-qPCR data are tabulated in Appendix 6.7. Greater than two-fold differences in *cir* transcript levels were observed in some *P. chabaudi* challenged mice, compared to the donor mouse, and are described below.

P. chabaudi challenge of mice immunized with the conserved CIR peptide resulted in lower levels of all *cir* transcripts measured, particularly those targeted by the primers U4 (PCHAS_040110, PCHAS_060140 and PCHAS_042030), Figure 48b and E1 (PCHAS_120060, PCHAS_104230 and PCHAS_000100), Figure 48d. These 2 - 8 fold decreases in *cir* detection, compared to the saline controls, were significant at the 95% confidence level, assessed by one-tailed Mann Whitney test.

The only other significant change in *cir* transcript levels was detected in $MSP1_{p21}$ immunized mice, which had a small (less than two-fold) decrease in the levels of *cir* genes targeted by primer pair E1, Figure 48d. In general, differences in *cir* transcript levels were variable between mice, meaning that few changes were significantly different from those observed in the saline controls. A clear trend was observed in *P. chabaudi* challenged mice immunized with the CIR sub-family A-F specific peptide, which had approximately 4-fold increased levels of *cir*s targeted by both sub-family U specific primers, Figure 48a and b.

Conversely, mice immunized with the CIR sub-family U specific peptide had increased levels of E1 *cir* transcripts, Figure 48d. This was also detected in mice immunized with the sub-family U recombinant protein, PCHAS_040110. In addition, the primer pair U4 detected lower levels of the equivalent *cir* gene in PCHAS_040110 immunized mice, compared to saline controls, Figure 48b.

The opposite trend appeared to occur in mice immunized with recombinant PCHAS_000100, as increased levels of the equivalent *cir* gene sub-set was detected, targeted by primer pair E1, Figure 48d. However, only one mouse could be analyzed for this result, thus further substantiation is required.

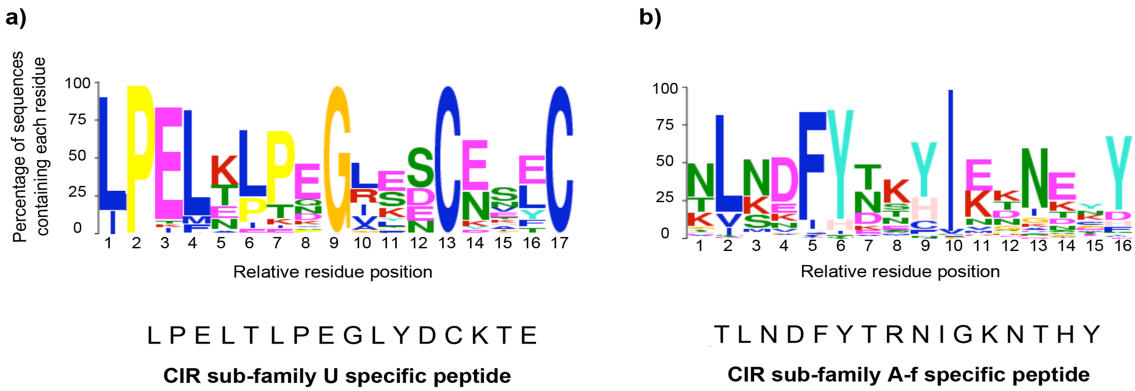


Figure 42: Design of CIR sub-family specific peptides.

The amino acid motifs corresponding to CIR sub-family U and CIR subfamilies A-F were found in 29.8% and 100.0% of the respective sub-family members. The proportion of CIR amino acid sequences containing each residue of the motifs are depicted as weblogo images (Crooks et al., 2004).

Peptides were designed using the sequences given above each amino acid motif. The sequences are given below each Weblogo image. The CIR sub-family U peptide was present in 37.0% of this sub-family, and the CIR subfamilies A-F peptide was present in 23.1% of these sub-families by BLAST search, allowing 5 mis-matches.

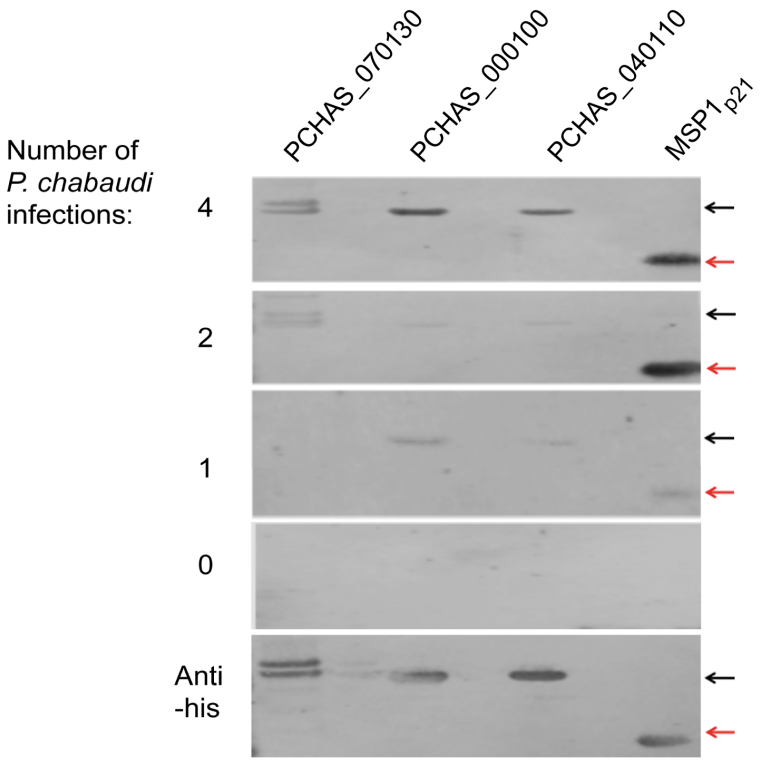


Figure 43: Recombinant CIR recognition by sera from a mouse experiencing one, two or four *P. chabaudi* AS infections.

500 ng of the recombinant CIR proteins: PCHAS_070130, PCHAS_000100, PCHAS_040110, and $MSP1_{p21}$ (Hensmann et al., 2004), were loaded onto a 12% SDS-PAGE gel and resolved under reducing conditions, where samples contained 100 mM DTT and the running buffer also contained 0.5 ml NuPAGE antioxidant.

Western blots were probed using serum from either a naïve BALB/c mouse (0 *P. chabaudi* infections), or a mouse that had experienced one, two or four *P. chabaudi* infections. In addition, antibodies recognizing the C-terminal His-tag present on all recombinant proteins were used as a positive control and as a loading control.

All blots show the recombinant CIR proteins at approximately 30 kDa, highlighted with black arrows, and recombinant $MSP1_{p21}$ at 14 kDa, highlighted with red arrows.

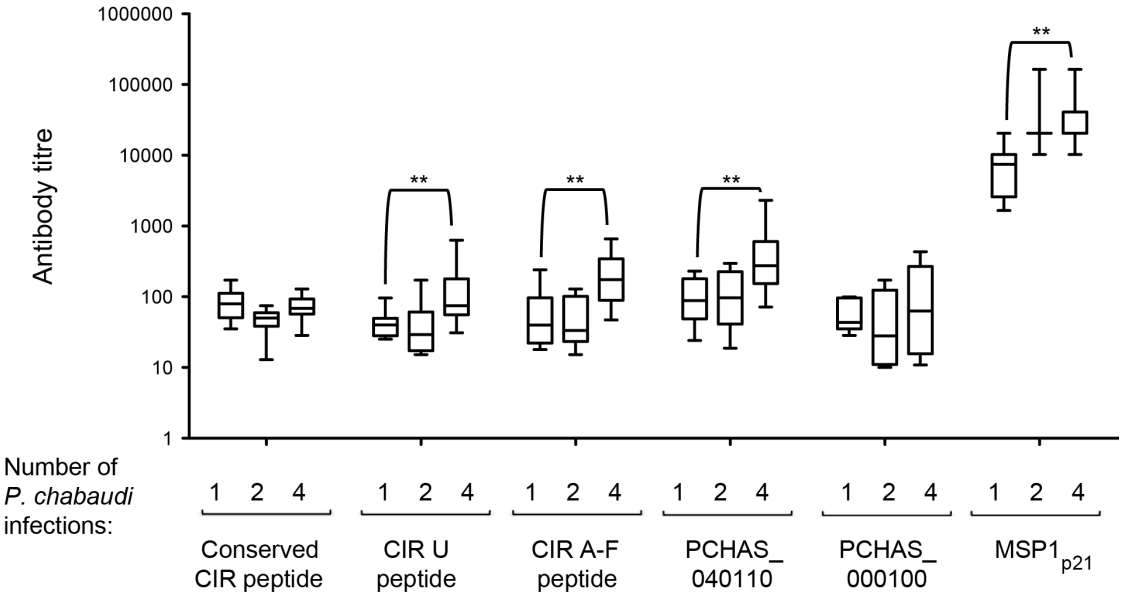


Figure 44: Titres of CIR-specific Abs, present in the sera of mice experiencing one, two or four *P. chabaudi* AS infections.

ELISAs were performed to measure the recognition of CIR and $MSP1_{p21}$ by Abs present in mouse sera. The recombinant CIR proteins PCHAS_000100, PCHAS_040110, or recombinant $MSP1_{p21}$, and the conserved CIR, CIR sub-family U and CIR sub-family A-F peptides were used as coating antigens.

The CIR proteins and peptides and $MSP1_{p21}$ were probed using serum from 10 BALB/c individuals that had experienced one, two or four *P. chabaudi* infections. Each protein was also probed with sera from a pool of 10 *P. chabaudi* naïve individuals, the absorbance of which was calculated as the 'baseline of non-specific Ab recognition'. This baseline was removed from the absorbance of *P. chabaudi* immune sera before Ab concentrations were approximated according to a known standard present on each ELISA plate: affinity purified mouse anti-PCHAS_000100 Abs.

These results are shown as box and whisker plots showing CIR or $MSP1_{p21}$ recognition by *P. chabaudi* immune sera from the 10 mice. The median and 5-95 percentile are shown, meaning that 90% of Ab titres are plotted [computed in Prsim v5.0 using the equations: 5*(N+1)/100 and 95*(N+1)/100)].

For each antigen, one-tailed Mann-Whitney tests were performed to determine whether Ab titres were significantly higher after four *P. chabaudi* infections, compared to one. Significance to the 99% confidence level is denoted by **.

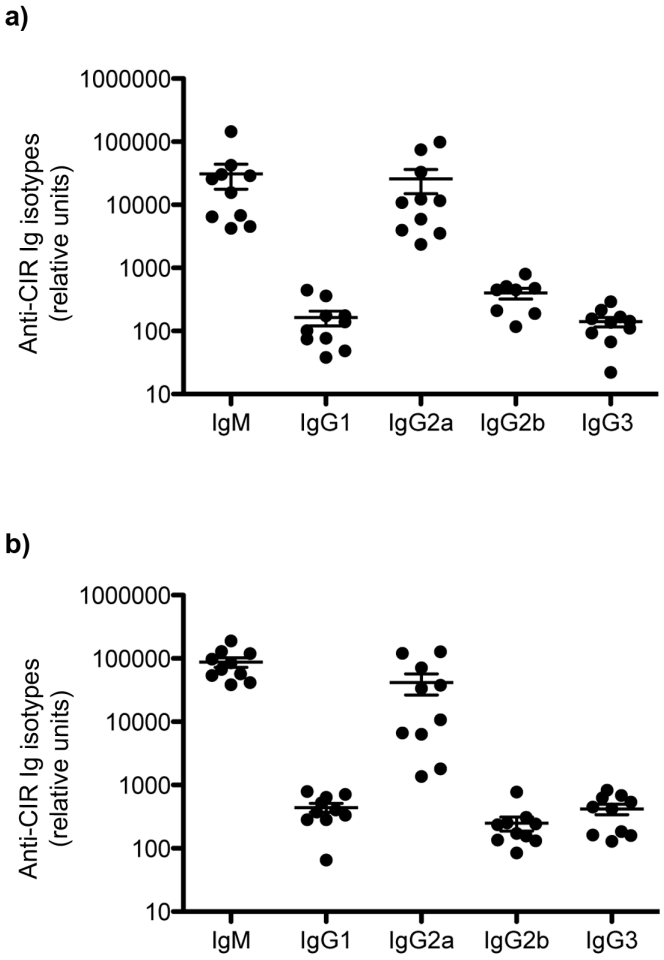


Figure 45: The isotype distribution of CIR-specific Abs present in sera from mice experiencing four *P. chabaudi* AS infections.

ELISAs were performed using the CIR sub-family U peptide, **a)**, and the CIR sub-family A-F peptide, **b)**, as coating antigens.

The isotype distribution was analyzed using immune sera from mice that had experienced four *P. chabaudi* infections, from the same 10 individuals used to estimate the concentration of total anti-CIR Ig, Figure 44. Each isotype concentration was estimated from a titration of purified mouse immunoglobulin for that isotype.

The concentration of CIR-specific Abs present in serum from each *P. chabaudi* immune mouse is plotted individually, each point representing a different mouse. Bars represent the geometric mean and 95% confidence intervals.

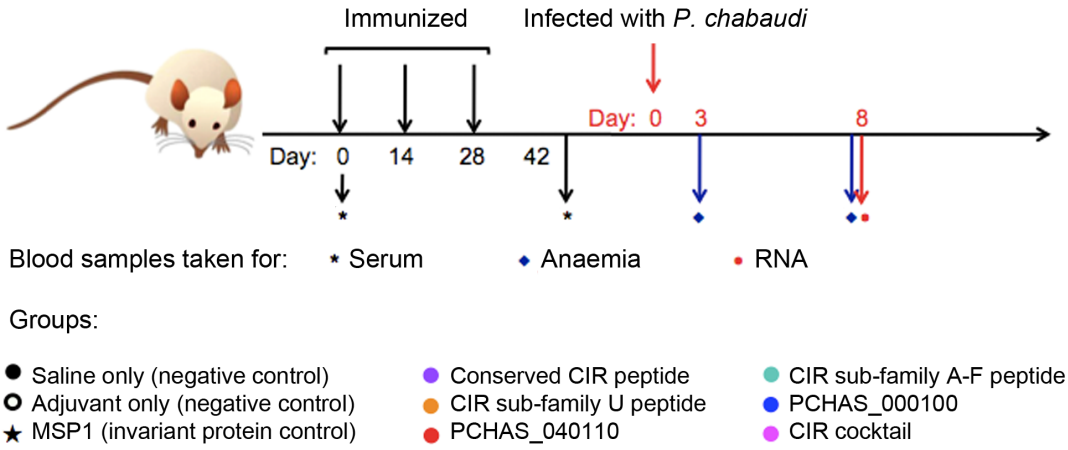
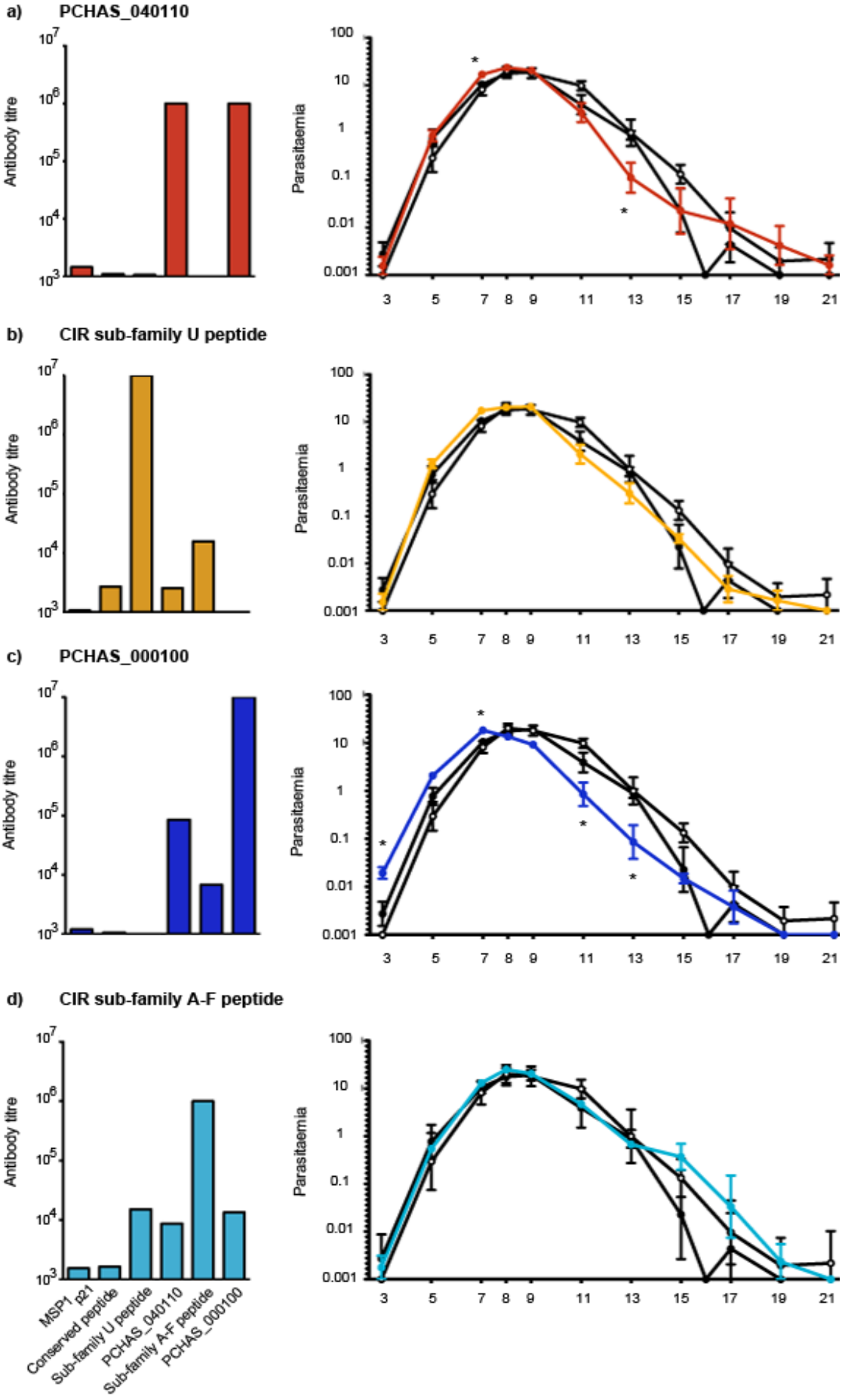


Figure 46: The experimental outline for CIR immunization and *P. chabaudi* challenge experiments.

Mice were immunized three times with 50 µg of CIR protein or peptide, $MSP1_{p21}$ or saline, in the presence of 12 µg Abisco adjuvant (Iscanova), according to the schedule shown in black. In addition a negative control of saline without adjuvant was performed to ensure no adjuvant-based effects were observed. The full list of groups is shown, along with the colour code used in all figures relating to this experiment.

Serum samples were taken prior to immunization, and 12 days after the final immunization to enable determination of induced Ab titres. 14 days after the final immunization, mice were infected with *P. chabaudi* AS, shown in red. Blood samples were taken for measurement of anaemia shortly after infection, and at the expected peak of parasitaemia, shown in blue. Blood samples were also taken at the expected peak of parasitaemia for analysis of *cir* transcription by RT-qPCR, shown in red.


a)
PCHAS_040110
b)
CIR sub-family U peptide
c)
PCHAS_000100
d)
CIR sub-family A-F peptide
Antibody titre
Parasitaemia
MSP1 p21
Conserved peptide
Sub-family U peptide
PCHAS_040110
Sub-family A-F peptide
PCHAS_000100

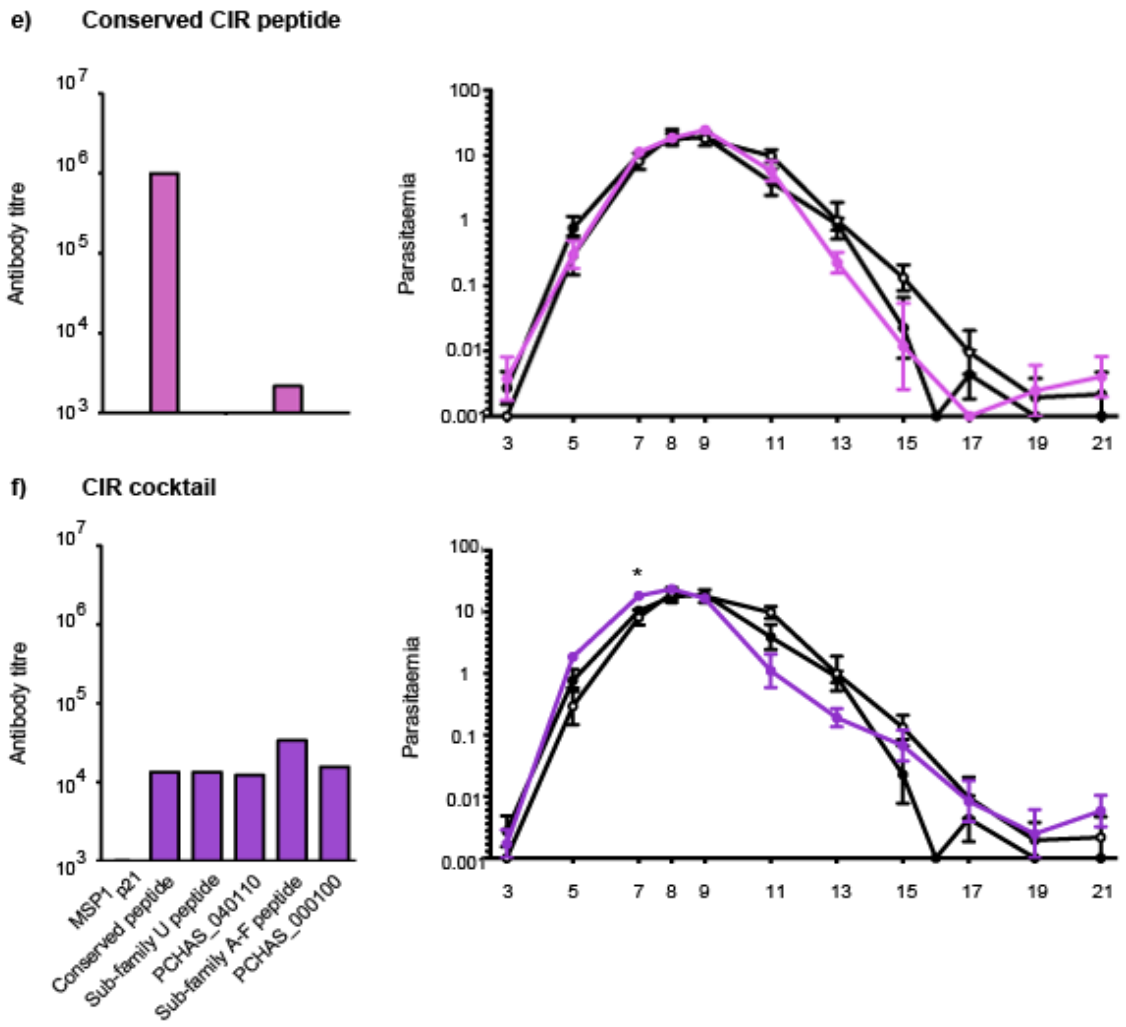


Figure 47: The titres of specific and cross-reactive Abs induced by immunization with CIR proteins.

Groups of mice were immunized with the following: the recombinant CIR protein PCHAS_040110, **a)**, CIR sub-family U peptide, **b)**, recombinant CIR PCHAS_000100, **c)**, the CIR sub-family A-F peptide, **d)**, conserved CIR peptide, **e)**, CIR cocktail, **f)**.

Left panels: Sera from all mice were compared before and after immunization, for their ability to recognize: recombinant $MSP1_{p21}$, the conserved CIR peptide, the CIR sub-family U and A-F specific peptides, the recombinant CIRs PCHAS_040110 and PCHAS_000100 (present in CIR sub-families U and E, respectively). Titre was defined as the dilution of immunized serum at which it reached the OD405 nm of serum from the same mouse prior to immunization, calculated using the mean of two technical replicates. Error bars represent the standard error of the mean.

Right panels: The courses of parasitaemia, are shown for mice infected with 10^3 *P. chabaudi* AS iRBC following immunization. All graphs show the courses of parasitaemia for the two negative control groups, mice immunized with saline only or saline in the presence of Abisco adjuvant (Iscanova). The geometric mean parasitaemia for each group of 5 mice is shown, excluding an individual that did not receive a full infectious dose in two groups. Error bars represent the standard error of the mean. One-tailed Mann-Whitney tests were performed to determine whether parasitaemias of CIR-immunized mice were significantly different from control mice. Significance to the 95% confidence level by comparison to both saline-immunized groups is denoted by *.

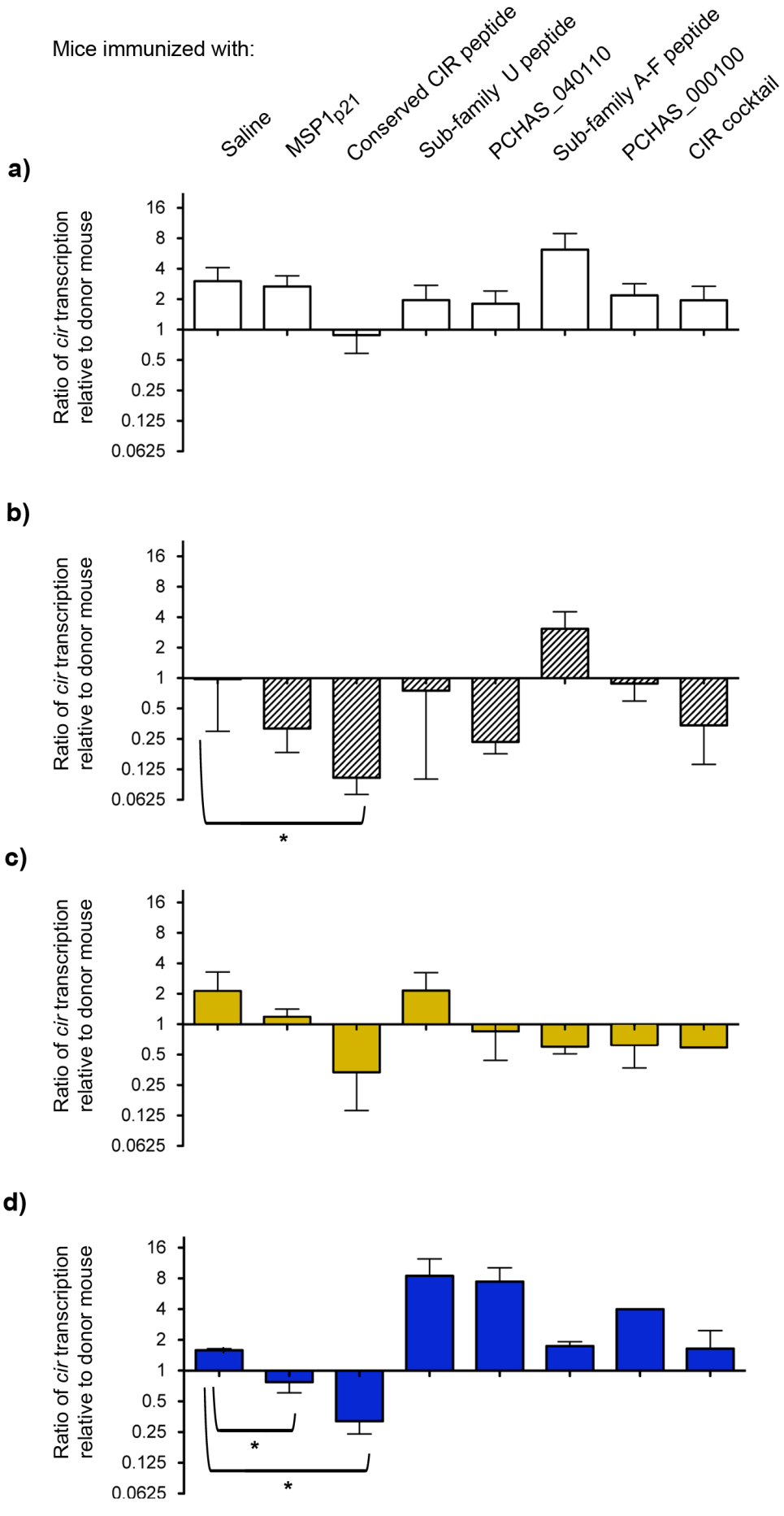

Mice immunized with:
Saline
MSP1p21
Conserved CIR peptide
Sub-family U peptide
PCHAS_040110
Sub-family A-F peptide
PCHAS_000100
CIR cocktail
a)
b)
c)
d)
Ratio of cir transcription relative to donor mouse
16
8
4
2
1
0.5
0.25
0.125
0.0625
*
*
*

Figure 48: The patterns of *cir* transcription observed in CIR-immunized mice challenged with *P. chabaudi* AS infection.

RT-qPCR was used to analyze 5 ng RNA from *P. chabaudi* infected blood samples taken on Day 8 of infection. The ratio (arbitrary units) of *cir* transcription normalized to the parasite reference gene, *beta tubulin*, was calculated for three immunized mice, relative to the *P. chabaudi* donor RAG2$^{-/-}$ mouse. For ease of view, these data are represented using a $\log_2$ scale, where a two-fold change in either direction is displayed at an equal scale. Ratios are not displayed as $\log_2$ values however.

Primer pairs used were: U1 (white bars), which target the *cir* genes PCHAS_040050 and PCHAS_001050 **a)**, U4 (hashed bars), which target the *cir* genes PCHAS_040110, PCHAS_060140 and PCHAS_042030 **b)**, C1 (yellow bars), which target the *cir* genes PCHAS_000140 and PCHAS_114750 **c)**, and E1 (blue bars), which target the *cir* genes PCHAS_120060, PCHAS_104230 and PCHAS_000100 **d)**.

The mean ratio of three mice analyzed per group is plotted. Mice were excluded if the SD between technical replicates exceeded 0.775 (80% confidence), attached in Appendix 6.7. Error bars represent the standard error of the mean.

One-tailed Mann-Whitney tests were performed to determine whether the ratios of *cir* transcripts detected in CIR-immunized mice, relative to the donor mouse, were significantly different from saline-immunized controls. Significance to the 95% confidence level by comparison to the group immunized with saline in the presence of adjuvant is denoted by *.

6.4 Discussion

In this chapter, efforts were made to determine whether CIR proteins were immunogenic during *P. chabaudi* infection, and the effect that induced immunity to CIRs could have upon subsequent infection. To accomplish these aims, CIR reagents were required. The generation of recombinant CIR proteins to represent each major CIR sub-family was described in chapter 3. In addition, a peptide was designed from a conserved CIR amino acid motif, found in the majority of CIR sequences, as described in chapter 5.

Antibodies to this peptide detected many features of CIR localization within iRBCs and merozoites which appeared similar to that described for the RIFINs of *P. falciparum* (Petter et al., 2007). The A and B-type RIFINs display different cellular localizations (Petter et al., 2007), which may be re-capitulated by the two major CIR sub-families. If this were truly the case, then members of CIR sub-family U would be predicted to localize at the iRBC surface like A-type RIFINs as they contain a similar insertion sequence (described in chapter 2.4). As such, members of the CIR U and A-F sub-families might experience different levels of immune pressure, and sub-family specific peptides were designed to determine whether this was the case. Briefly, the CIR sub-family U specific peptide contained similar features to the RIFIN A-type insertion sequence, and was present within the recombinant CIR PCHAS_040110; whilst the CIR sub-family A-F peptide corresponded to a similar sequence in the recombinant CIR PCHAS_000100.

All CIR proteins and peptides were recognized by the immune system after *P. chabaudi* infection; with the exception of the conserved CIR peptide, which is found within the hydrophobic core of CIR proteins (discussed in chapter 5). Titres of Abs recognizing CIRs were approximately 100 fold lower than those recognizing recombinant $MSP1_{p21}$. Even after four *P. chabaudi* infections, the majority of Abs recognizing the CIR sub-family specific peptides comprised IgM and IgG2a isotypes. Immunization with CIR proteins and peptides was successful in that the sera of all mice contained Abs recognizing their cognate antigen. CIR immunization did not affect parasitaemia upon *P. chabaudi* challenge, apart from a slight increase in certain immunized groups on day 7 of infection. The exception was for the mice immunized with PCHAS_000100, which experienced an earlier peak of infection, and earlier parasite clearance than all other groups. Infected RBCs from these mice did not display any difference in *cir* transcript

levels, relative to the donor mouse, compared to the saline controls. Infected RBCs from mice immunized with PCHAS_040110 however, had significantly lower levels of the corresponding *cir* gene or *cir*s closely related to it.

The titre of Abs present in the sera of *P. chabaudi* immune mice recognising CIRs was lower and displayed larger variation than Abs recognizing $MSP1_{p21.}$ MSP1 is expressed by all parasites and is therefore probably more exposed to the immune system than individual CIR proteins, which are predicted to be differentially expressed. This difference in protein expression probably led to the stronger Ab response detected against $MSP1_{p21}$ than any of the CIRs tested. It has been hypothesized that the majority of CIR proteins are expressed at levels below the threshold sufficient for induction of successful immune responses, thus providing a smoke-screen behind which parasites may continue to replicate (Cunningham et al., 2009). The transcription of many *cir* genes during *P. chabaudi* infection supports this idea (described in chapter 3.3.2), as do the low level Ab responses observed to CIRs even after four *P. chabaudi* infections. The highest Ab responses were to the CIR sub-family A-F specific peptide and PCHAS_000100 (sub-family E), followed by the CIR sub-family U specific peptide and the corresponding recombinant CIR PCHAS_040110.

The Ab isotypes recognizing both sub-family specific CIR peptides were predominantly IgM and IgG2a, in the sera of mice which had experienced four *P. chabaudi* AS infections. IgM are usually the first Abs which appear during infection and do not undergo affinity maturation or somatic hypermutation to improve the specificity of antigen binding (Janeway, 2007). Nonetheless, the pentameric structure of IgM ensures high avidity binding to antigen, enables efficient complement fixation and can promote the lysis and / or agglutionation of pathogens (Janeway, 2007). The high levels of IgM Abs detected in response to CIR proteins have two plausible explanations. First, that few CIR-specific memory B cell responses were induced following *P. chabaudi* infection thus IgM provided an early defence at each infection. Or second, that memory did develop against CIRs, but that fewer CIR-specific B cells underwent maturation into IgG2a production, thus memory responses remained producing IgM.

By contrast, IgG2a Abs are only produced after affinity maturation, somatic hypermutation and class switching, and are thus characteristic of a memory response (Janeway, 2007). As IgG2a was the predominant IgG isotype detected for CIRs after four *P. chabaudi* infections, this likely represents a memory response, as also detected for $MSP1_{p21}$ in CB57Bl/6 mice [where the the equivalent isotype is IgG2c (Jouvin-

Marche et al., 1989, Morgado et al., 1989)]. This Ab isotype has been shown to induce complement activation and bind Fcγ receptors more effectively than other IgG isotypes (Nimmerjahn and Ravetch, 2006, Nimmerjahn and Ravetch, 2005, Nimmerjahn et al., 2005). Passive transfer of IgG2a Abs from *Plasmodium*-immune mice to naïves have also been shown to provide more protection than other isotypes, for example; delaying the onset of parasitaemia for *P. yoelii* infection (White et al., 1991), and inhibiting RBC re-invasion for *P. chabaudi* AJ (Cavinato et al., 2001), compared to IgG1 Abs. The contribution of IgG2a Abs recognizing CIR proteins to these effects is as yet unknown.

The immune recognition of other PIRs has only been investigated in two studies to date, focussing on the VIRs. Fernandez-Becerra and colleagues found that approximately three quarters of *P. vivax* infected patients had generated Abs recognizing at least one of the 22 VIR proteins analyzed, showing that VIR proteins were immunogenic in natural infection (Fernandez-Becerra et al., 2005). This work was extended by Oliveira and colleagues, who surveyed the reactivity of 200 patients' sera against seven recombinant VIRs (Oliveira et al., 2006). Approximately half of all *P. vivax* patients had a measurable IgG or IgM response to at least one of the seven VIRs, measured by ELISA. By contrast, 90.5% of individuals had measurable IgG responses against $MSP1_{19}$. Similarly, titres of Abs in the sera of *P. chabaudi* infected mice were consistently higher against $MSP1_{p21}$ than CIRs.

IgG and IgM Abs recognizing at least one of the 7 recombinant VIRs were found respectively in the sera of 26% and 29.6% of *P. vivax* infected individuals. The high proportion of IgM recognizing VIRs appears similar to what was detected for CIRs, however, a definitive comparison cannot be made as no distinction was made between the number of *P. vivax* infections experienced by the patients (Oliveira et al., 2006).

Unlike Ab responses to CIR proteins, which increased after each *P. chabaudi* infection, the proportions of patients making a positive IgG response to VIR did not increase with the number of *P. vivax* infections experienced (Oliveira et al., 2006). There was no bias in the isotype of IgG Abs in patient sera recognizing recombinant VIR proteins from each sub-family (Oliveira et al., 2006). This suggests that, unlike for the CIRs, anti-VIR Ab responses were not dominated by a particular IgG isotype, although again, comparison is difficult owing to the varying number of *P. vivax* infections experienced by the different patients surveyed.

Dissection of anti-PIR immunity is complicated in the field, due to variation between individuals and the infecting parasites. Equally, the presence of Abs recognizing PIR proteins does not inform whether they have any role in protection against either infection or disease. In fact many 'blocking Abs' can be induced in response to particular epitopes of *Plasmodium* proteins, which prevent effector Abs from reaching their target [for example, (Uthaipibull et al., 2001)]. *P. chabaudi* challenge of CIR-immunized mice was thus carried out to determine whether immunity generated against CIR proteins could play a role in the outcome of the infection. Immunization to induce high titres of anti-CIR Abs was carried out in the presence of Abisco adjuvant, which was derived from saponin. Adjuvants influence the titre of induced Abs, their isotypes and avidities, and can mediate other effects (Oda et al., 2000). Saponin based adjuvants are useful as they can stimulate cell mediated immunity in addition to enhancing Ab production (Oda et al., 2000). Due to time constraints, the isotype distribution of Abs induced by CIR-immunization was not investigated, and should be ascertained. However, it is known that immunization in the presence of saponin based adjuvants often induces high titres of IgG1 and IgG2a Abs [for example (Fleck et al., 2006), http://www.isconova.com/products/abisco-research-reagent/abisco-100.aspx]. This was likely to also be the case for the CIR-immunized mice.

All immunized mice made Ab responses to the antigen they were immunized with. Abs present in the sera of mice immunized with sub-family specific CIR peptides also recognized the corresponding recombinant CIR protein, probably due to recognition of the equivalent peptide within the recombinant CIRs by these Abs. The two recombinant CIR proteins, whilst representing different major sub-families, were likely to share conformational determinants, which would induce cross-reactive Abs after immunization. This may explain why sera from mice immunized with one recombinant CIR contained Abs that could also recognize the other one.

Upon challenge of all other CIR-immunized mice there were few differences in parasitaemia compared with saline immunized controls. This could have been due to several reasons. The first, that induced Ab titres were too low to have an effect, may have applied to mice immunized with the CIR cocktail, as the sera of these mice contained very low titres of Abs against all CIR antigens. The second explanation could be that the CIR targets were not exposed to the immune system for induced anti-CIR Abs to have an effect. This is the most likely explanation for mice immunized with the conserved CIR peptide, as the hydrophobic core of CIR proteins was probably

inaccessible to these Abs. The relevant CIRs in other immunized mice may also have remained within the iRBC or were not expressed by many parasites in the infecting population. In the latter case, iRBCs expressing CIRs not recognized by the induced Abs would have a selective growth advantage in the immunized mice.

Anti-CIR Abs induced by immunization could only contribute to protective effects if their protein target were exposed during *P. chabaudi* infection. Some CIR proteins were detected near to or at the iRBC surface (described in chapter 5), which may correlate to the members of CIR sub-family U, if the CIR sub-families do display similar localizations to those of the A- and B- type RIFINs. Thus, mice immunized with the CIR sub-family U peptide and protein PCHAS_040110 were expected to be protected better than mice immunized with the CIR sub-family A-F peptide or PCHAS_000100. Surprisingly however, the reverse was observed, and mice immunized with PCHAS_000100 experienced earlier parasite clearance than all other groups. The reason for this is unclear. It could be that PCHAS_000100 was simply more representative of native CIRs than PCHAS_040110, that the higher titres of Abs recognizing this protein were able to cross-react with other members of the CIR repertoire, that PCHAS_000100 was the dominant CIR protein expressed during the *P. chabaudi* challenge infection, or a combination of these.

Another interesting possibility comes from the observation in chapter 5, that CIRs were located at the apical end of merozoites. Cavinato and colleagues have suggested that Abs present in the sera of *P. chabaudi* AJ 'hyper-immune' mice could predominantly reduce parasite numbers at the RBC re-invasion stage rather than, for example, by opsonization of iRBCs (Cavinato et al., 2001). This is likely to act upon merozoite proteins such as MSP1 and AMA1, but it is possible that CIRs are also affected by such mechanisms. However, as the effect of Abs recognizing PCHAS_000100 was only apparent after day 7 of *P. chabaudi* challenge, it is likely that they did not act to inhibit RBC invasion. Instead, other mechanisms may have contributed to parasite clearance in PCHAS_000100 immunized mice, which may have involved Ab recognition of merozoites rather than iRBCs.

A trend for increased parasitaemia was observed on day 7 of infection in four of the CIR immunized groups. Peripheral parasitaemia was measured at 9pm in the challenged mice (housed under normal light conditions, where schizogony would be expected to occur at midnight). It is possible that Abs induced by CIR-immunization were able to prevent sequestration of late-trophozoite iRBCs at this time, causing an efflux of iRBC

in the periphery, compared to control mice where the iRBCs were able to sequester in the deep vasculature. The VIR proteins have recently been implicated in cyto-adhesion of some *P. vivax* iRBCs, albeit in a preliminary experiment where not all anti-VIR anti-sera produced equivalent effects (Carvalho et al., 2010). If the ability of VIR proteins to mediate cyto-adhesion is substantiated, other PIR proteins could also play a role in sequestration. Then, the above mechanism could explain the higher parasitaemia observed in mice immunized with conserved CIRs or those predicted to be at the iRBC surface.

No substantial differences between the $MSP1_{p21}$ and saline immunized mice were observed, confirming similar results obtained during *P. chabaudi* challenge of mice immunized with the recombinant C-terminus of MSP1 (O'Dea et al., 1996, Hensmann et al., 2004). This can be explained by the fact that low numbers of CD4 T-cells recognize the C-terminal region of MSP1 (Quin and Langhorne, 2001). During infection, CD4 T-cells recognizing epitopes in the central region of MSP1 provide necessary help to mount an effective Ab response against the C-terminus (Quin and Langhorne, 2001). However, after immunization with only the C-terminal region, such help may be lacking. Low proliferative T cell responses have also been observed in relation to VIR proteins (Oliveira 2006). This may indicate that PIR proteins, like $MSP1_{p19}$ and its equivalents, fail to induce effective CD4 help for Ab production, which would also explain the high levels of IgM observed in response to CIR proteins, after four *P. chabaudi* AS infections, Figure 45.

Since parasitaemia was expected to be too crude a measure to detect changes in the parasite response to anti-CIR immunity, parasite transcription was measured at the expected peak of infection. Whilst the majority of parasites displayed no substantial changes in *cir* expression compared to the donor mouse, there were some differences. For example, significantly lower levels of *cir* sub-family E and U transcripts were detected in mice immunized with the conserved CIR peptide, than in control mice, immunized with saline. Non-significant, but substantial changes were also observed upon *P. chabaudi* challenge of mice immunized with sub-family specific reagents. The detection of *cir* sub-family U transcripts increased after immunization with the opposite CIR sub-family A-F specific peptide. The converse, increased levels of *cir* transcripts belonging to sub-family E, was observed in mice immunized with the sub-family U specific peptide and the recombinant sub-family U CIR protein PCHAS_040110. In

addition, iRBCs isolated from mice immunized with PCHAS_040110, had reduced levels of the corresponding transcript.

These data suggest that in the face of a CIR protein or sub-family specific immune response, iRBCs expressing less of those *cir*s (and instead transcribing others) were able to dominate the *P. chabaudi* infection. This could have been due to immune-mediated destruction of iRBCs expressing the corresponding CIR protein(s), and subsequent expansion of an iRBC population expressing different CIRs. Alternatively, switching of *cir* / CIR expression could have allowed iRBCs to evade specific immunity. This phenomenon was characterized for the VSAs of *P. knowlesi* and is also known to occur in *P. chabaudi*, permitting establishment of chronic infection (Brown and Brown, 1965, McLean et al., 1982, McLean et al., 1986). Whilst no change was observed in the majority of iRBCs, this may reflect the limited detection of *cir* family members by RT-qPCR and *cir* switching events may have occurred without detection.

It is also possible that other related *cir* gene transcripts may have increased, which were not detected by RT-qPCR. The limited number of *cir*s that could be analyzed by this methodology means that whole transcriptome sequencing of the parasites in *P. chabaudi* challenged mice may be more informative than RT-qPCR. If the sensitivity of RNA sequencing methods were improved (as discussed in chapter 3.4), this would allow detection of the switch in dominant *cir* transcript between the infections of donor and challenge mice. Similarly, proteomic techniques could be used for the detection of dominant CIR proteins.

The experiments presented here indicate that not only are CIR proteins recognized by the immune response following *P. chabaudi* infection, but immunization with CIR can affect both the parasitaemia during a *P. chabaudi* challenge and also the parasite transcriptional response, whether by immune selection of different parasites or by active alteration of *cir* transcription in the face of a specific Ab response. Future experiments should be carried out to elucidate the true effect of CIR immunization, using larger group sizes and a lower infectious dose of *P. chabaudi* to enable clearer observation of the effects of anti-CIR Abs. In addition, it is feasible that immunization with CIR proteins could reduce the proportion of iRBC in the population that are associated with parasite virulence, without affecting total parasitaemia. To investigate this possibility, measures of pathology such as weight loss, reduction in temperature, lethargy, and anaemia in the mice should be implemented.

The recent development of technologies to permit effective *P. chabaudi* transfection (Spence et al., 2011) has allowed the generation of luminescent *P. chabaudi* in our laboratory. These parasites could be used to indicate sites of *P. chabaudi* sequestration by whole body imaging and subsequent dissection of organs. If CIR proteins are involved in sequestration, CIR-immunized mice may display differences in the parasite distribution during challenge infection.

Another application of *P. chabaudi* transfection (Spence et al., 2011) could be to manipulate CIR expression in these parasites. Unfortunately, given the large family size, *cir* 'knock-outs' would not be practical to attempt, meaning that subsequent complementation with a single *cir* would not be possible. Thus, analysis of a single *cir* in isolation would probably be infeasible. An alternative strategy could be to over-express a particular *cir* gene, such that all transfectants would display the same CIR at the iRBC surface, as discussed in chapter 5. Immunizing mice with that CIR protein may then permit the detection of specific immune responses more clearly.

Chapter 7: Final perspectives

The work presented in this thesis has sought to characterize the *cir* multi-gene family of *P. chabaudi*, by analysis of their sequences, transcription, protein structure, cellular localization and immune recognition. 198 *cir* genes were identified in the *P. chabaudi* genome, 86% of which clustered to form two major sub-families and six minor ones, on the basis of sequence similarity. Reagents were designed to analyze *cir* gene and CIR protein expression during the blood stages of *P. chabaudi* infection. Transcripts belonging to all *cir* sub-families were detected at low levels by RT-qPCR and RNA sequencing. Each RNA sequencing sample contained a dominant *cir* transcript, both of which were members of *cir* sub-family U. CIR localization in blood stage parasites was determined by confocal microscopy and flow cytometry, detecting CIRs at the apical end of merozoites and located within the PV of iRBCs. In addition, CIRs were also found at the infected erythrocyte surface of some mature trophozoites. After *P. chabaudi* infection, low titres of Abs were detected, which were predominantly IgG2a and IgM isotypes. Immunization with CIRs mostly did not influence subsequent *P. chabaudi* challenge either by altering parasitaemia or the levels of *cir* transcripts. However, one of the recombinant proteins resulted in significantly earlier parasite clearance, and whilst the other did not, the corresponding transcript was reduced.

Many similarities to other members of the PIR super-family were detected in the CIR repertoire, in their amino acid sequence features and the identification of sub-families containing similar sequences [chapter 2, (Del Portillo et al., 2001, Janssen et al., 2002, Janssen et al., 2004, Merino et al., 2006, Fonager et al., 2007)]. Other features shared by CIRs and other PIRs included the detection of many transcripts during the blood stages of infection [chapter 3, (Bozdech et al., 2008, Cunningham et al., 2009)], localization of proteins near to or at the surface of trophozoite stage iRBCs [chapter 5, (Del Portillo et al., 2001, Janssen et al., 2004, Cunningham et al., 2005, Fernandez-Becerra et al., 2005, Petter et al., 2007, Niang et al., 2009, Joergensen et al., 2010, Khattab and Meri, 2011)], and relatively low titres of Abs recognizing these proteins detected in the sera after *Plasmodium* infection [chapter 6, (Oliveira 2006)].

The PIR super-family had previously been proposed to include the RIFINs of *P. falciparum*, however, evidence to support this beyond similarities of the intron and amino acid sequences was lacking (Janssen et al., 2004). Several features observed in

common for the RIFIN and CIR repertoires, presented in this thesis, suggest that indeed the PIR and RIFIN proteins may form part of a *Plasmodium*-wide multi-gene family. Notably, both CIR and RIFIN repertoires cluster to form two major sequence sub-families, one of which is predicted to have undergone functional divergence, whose members contain an insertion containing two conserved cysteine residues (Petter et al., 2007, Joannin et al., 2008). The close position of the cysteine residues was similar between the sub-family A RIFIN and CIR sub-family U amino acid motifs. Cysteine residues often participate in the formation of disulphide bridges [reviewed by (Sevier and Kaiser, 2002)], which suggests that the differential cysteine compositon of CIR and RIFIN sub-families could affect their tertiary structure. The structure of both PIR and RIFIN proteins was predicted to comprise alpha helices of 12-20 amino acids, separated by coiled-coil regions (Janssen et al., 2004). In this thesis, the first experimental evidence for PIR protein structure was presented, using the recombinant CIR protein PCHAS_000100. Circular dichroism detected 44% alpha helices within this protein, a measurement which falls within the range predicted by Janssen and colleagues (Janssen et al., 2004), and also suggests that PCHAS_000100 was folded similarly to native CIR proteins.

The majority of PIR proteins are predicted to contain at least one TM domain, a feature which has led to predictions that the C-terminus of most PIR proteins is intra-cellular, whilst the N-terminus is exposed at the iRBC surface (Del Portillo et al., 2001, Del Portillo et al., 2004). Pain and colleagues have proposed the existence of a PIR protein domain, which may be repeated up to three times beyond the TM domain (Pain et al., 2008), and is likely to also be present within the CIR proteins. Informatic predictions have suggested that B-type RIFINs may contain two TM domains, whilst sub-type A proteins contain one (Bultrini et al., 2009). Modelling the structure of an A-type RIFIN from primary amino acid sequence has indicated the presence of an Armadillo fold (Bultrini et al., 2009), which is characterised by alpha-helices arranged to form a wide cleft with an extensive solvent-accessible surface and typically able to bind large substrates [reviewed by: (Coates, 2003)]. Further experimental analyses are required to verify this prediction, which may gain insight into potential binding partners.

Members of the two RIFIN sub-families have been shown to display different cellular localizations (Petter et al., 2007). Whilst this has not shown for the two major CIR sub-families, Abs recognizing a broad range of CIRs reproduced both patterns observed for the A and B-type RIFINs: within the iRBC cytoplasm and within the PV [chapter 5,

(Petter et al., 2007)]. If a similar distinction in protein localization were confirmed for the two CIR sub-families, this would provide strong support to the inclusion of PIRs and RIFINs in one super-family.

The dedication of a large proportion of the *Plasmodium* genome to accommodate multi-gene families indicates that their encoded proteins may be important for parasite survival. Similarities identified between the PIRs and RIFINs have suggested that these proteins are analogous, and may form a ubiquitous multi-gene family, present in all *Plasmodium* species sequenced to date. If this is the case, STEVOR and *Pf*MC-2TM proteins may also be included due to their shared sequence characteristics (described in chapter 1.8). Thus, it is important to uncover the function(s) of these proteins, which would aid their classification, and fundamentally aid understanding of malaria parasites.

The detection of CIR proteins exposed at the surface of some iRBCs in chapter 5, provides the first piece of evidence that CIRs may interact with the host immune response. YIR proteins have also been detected in this location, and *yir* transcription shown to change substantially during a secondary *P. yoelii* infection of immuno-competent mice, but not in immuno-compromised animals (Cunningham et al., 2005). It has been hypothesized that the majority of PIR proteins are expressed at levels below the threshold sufficient for induction of successful immune responses, thus providing a smoke-screen behind which parasites may continue to replicate (Cunningham et al., 2009). The transcription of many *cir* genes during *P. chabaudi* infection supports this idea (chapter 3), as do the low level Ab responses observed to CIRs even after four *P. chabaudi* infections (chapter 6).

Any CIRs exposed at the iRBC surface could potentially mediate antigenic variation. *P. chabaudi* is known to express variant antigens during the late trophozoite stages of development in the bloodstream (Brannan et al., 1994), and to undergo immune evasion (McLean et al., 1982, McLean et al., 1986). Antigenic variation has also been shown to correlate with the ability of *P. chabaudi* iRBCs to sequester and to maintain chronic infection (Gilks et al., 1990). Whilst the antigens responsible have not been definitively characterized, early experiments were able to clone a *P. chabaudi* gene, of which the expressed fusion protein retained reactivity with variant specific anti-sera and induced some protection during an immunization and challenge experiment (Phillips et al., 1997). The whole gene product was 597 amino acids (Phillips et al., 1997), within the size range of the *cir* repertoire.

Evidence that may support a role for *cir* genes in antigenic variation was observed during the *P. chabaudi* challenge experiment, presented in chapter 6. Mice immunized with CIR sub-family specific reagents had increased levels of *cir* transcripts belonging to the other major sub-family, and some decrease of the corresponding or similar *cir* transcripts, during a *P. chabaudi* challenge infection. It is likely that immune selection against iRBCs expressing the equivalent proteins favoured the growth of parasites expressing other CIRs, recapitulating the differences in *P. chabaudi* antigens previously observed between primary and recrusdescent peaks of infection (McLean et al., 1982, McLean et al., 1986).

Another phenomenon observed during the *P. chabaudi* challenge experiment was a trend for increased parasitaemia on day 7 of infection in certain CIR-immunized groups. As parasitaemia was measured at 9pm, when parasites would be predominantly late trophozoite stage, it is possible that Abs induced by CIR-immunization were able to prevent sequestration at this time. This would cause an efflux of iRBCs into the peripheral blood, compared to control mice where the iRBCs were able to sequester in the deep vasculature. It is possible that PIR proteins could mediate cyto-adhesion in the deep vasculature in regions where less rapid blood flow occurs, as non *P. falciparum* species are known to sequester to some extent (Mota et al., 2000, Franke-Fayard et al., 2005, Carvalho et al., 2010, Gilks et al., 1990). *P. vivax* iRBCs have recently been shown to bind endothelium with an equivalent strength to *P. falciparum* iRBCs, although the attachment process was slower (Carvalho et al., 2010). Preliminary data has implicated the VIR proteins in this process (Carvalho et al., 2010), although this awaits confirmation.

Predictions based upon features of known adhesive proteins (Sachdeva et al., 2005), when applied to *Plasmodium* have suggested that PIR and RIFIN proteins could be potential parasite adhesins (Ansari et al., 2008). Besides sequestration, another role for adhesive proteins is during attachment and invasion of RBCs. The detection of CIR proteins at the apical end of merozoites (chapter 5), similar to A-type RIFINs and STEVORs (Blythe et al., 2008), suggests that these proteins mediate functions in several stages of parasite development. STEVOR proteins additionally appear to coat the merozoite surface (Khattab and Meri, 2011). Anti-STEVOR Abs have been postulated to inhibit RBC invasion (Garcia et al., 2005), however, to date the evidence is not clear. Different studies have shown anti-STEVOR Abs and peptides either to inhibit RBC invasion or to have no effect, according to the concentration used (Garcia

et al., 2005, Khattab et al., 2008). If STEVOR proteins are truly involved in RBC invasion, it has been postulated that their biological role could involve the deflection of specific immunity away from essential invasion machinery (Khattab et al., 2008).

Potentially, the PIR proteins could also play a role in RBC invasion process, and in fact approximately 30 *yir* genes were found to be up-regulated during *P. yoelii* challenge of mice immunized with merozoite surface protein 8 [MSP8, (Shi et al., 2005)]. MSP8 immunization is believed to protect mice from lethal *P. yoelii* 17XL infection by restricting parasite invasion to reticulocytes (Shi et al., 2005). Thus PIR proteins could conceivably act as the required alternative RBC invasion ligands or could shield the crucial ligands from the immune response.

Although not investigated here, CIR proteins could also play a role in other stages of the *P. chabaudi* life-cycle. Transcription or protein expression has been detected for PIRs and RIFINs in gametocytes, sporozoites and mosquito stages, (Florens et al., 2002, Hall et al., 2005, Petter et al., 2008, Wang et al., 2010). Their function(s) in these stages may involve evasion of mosquito immunity for ookinetes and innate mammalian immunity for sporozoites, or adhesion of oocysts. Insects and innate vertebrate responses are known to share features of complement and immune signalling (the insect pathways Toll, Imd and Hop, are equivalent to the mammalian TLR, TNFα and JAK/STAT pathways [reviewed by (Vilmos and Kurucz, 1998, Pinheiro and Ellar, 2006)]. In addition, a role for PIR proteins during invasion of the mosquito mid-gut and mammalian hepatocytes is also possible, if also true within the blood stages of the lifecycle. These possibilities are speculative and await evaluation.

The work presented in this thesis has raised several interesting lines of enquiry regarding the function of CIR proteins during the blood stages of *P. chabaudi* infection. A crucial question is whether proteins belonging to the two major CIR sub-families display differential sub-cellular localization, as observed for the RIFINs. Once answered, this will enable predictions about the role of CIRs during *P. chabaudi* infection, may illuminate results of the *P. chabaudi* challenge experiment and may strengthen the similarities observed between CIRs and RIFINs. The recombinant CIR proteins or sub-family specific peptides may be used to produce polyclonal Abs for this purpose.

Tools generated here may also be used to investigate the role of CIR proteins in cyto-adhesion, RBC invasion and antigenic variation during *P. chabaudi* infection. For

example, RNA sequencing could be used in combination with proteomic approaches, providing a powerful means of analysis for the whole *cir* repertoire. The sensitive nature of RT-qPCR could enable detection of certain *cir* transcripts in individual iRBCs, as shown previously for the *yir*s and *vir*s (Fernandez-Becerra et al., 2005, Cunningham et al., 2009).

If the PIR and RIFIN families are related, it is possible that the mechanisms regulating their expression are similar, and may incorporate features of *var* gene regulation (Howitt et al., 2009). *Plasmodium* VSAs often appear to be down-regulated during *in vitro* culture (Dzikowski and Deitsch, 2008, Howitt et al., 2009), thus functional analyses of VSA proteins would be best carried out *in vivo*. Investigations of VSA regulation in *P. falciparum* have focused on the *var* genes, which appear to require a substantial epigenetic contribution, as discussed in chapter 1.8.3. However, analyses of the regulation and functions of *Plasmodium* VSAs should also prioritize the PIR super-family, which may include the RIFINs, given their ubiquitous nature and likely importance for parasite survival.

VSAs encoded by *Plasmodium* multi-gene families are likely to mediate antigenic variation, in addition to other possible functions. Whilst antigenic variation has long been assumed to be responsible for chronic infection and the evasion of host immunity, evidence has only been generated recently, for the intestinal parasite *Giardia lamblia*. In this parasite, disruption of the RNA interference pathway, which regulates VSP expression (Prucca et al., 2008), resulted in simultaneous expression of all VSPs (Rivero et al., 2010). Immunization with purified VSPs from these parasites completely protected gerbils from *G. lamblia* infection, which may have applications for vaccine development. Similar disruption of antigenic variation in *Plasmodium* may also be possible if the mechanisms underlying VSA regulation were understood, which could have important applications to malaria vaccine design strategies.

Clearly there is still a great deal to learn about the large multi-gene families present in all *Plasmodium* species, particularly with respect to their possible host-interactions such as sequestration and immune evasion. These investigations are most accessible in rodent models, following which mechanisms can be confirmed by patient studies. The *cir* multi-gene family of *P. chabaudi* may provide an ideal model system for such research.

References

ABDEL-LATIF, M. S., DIETZ, K., ISSIFOU, S., KREMSNER, P. G. & KLINKERT, M. Q. (2003) Antibodies to *Plasmodium falciparum* RIFIN Proteins Are Associated with Rapid Parasite Clearance and Asymptomatic Infections. *Infection and Immunity,* 71**,** 6229-6233.

ABDEL-LATIF, M. S., KHATTAB, A., LINDENTHAL, C., KREMSNER, P. G. & KLINKERT, M. Q. (2002) Recognition of variant RIFIN antigens by human antibodies induced during natural *Plasmodium falciparum* infections. *Infection and Immunity,* 70**,** 7013-21.

ABHIMAN, S. & SONNHAMMER, E. L. L. (2005) Large-Scale Prediction of Function Shift in Protein Families with a Focus on Enzymatic Function. *PROTEINS: Structure, Function, and Bioinformatics,* 60**,** 758-768.

ABRAMOFF, M. D., MAGELHAES, P. J. & RAM, S. J. (2004) Image processing with imageJ. *Biophotonics International,* 11**,** 36-41.

ACHTMAN, A. H., KHAN, M., MACLENNAN, I. C. M. & LANGHORNE, J. (2003) *Plasmodium chabaudi chabaudi* infection in mice induces strong B cell responses and striking but temporary changes in splenic cell distribution. *Journal of Immunology,* 171**,** 317-324.

AIKAWA, M., UDEINYA, I. J. & RABBEGE, J. (1985) Structural alteration of the membrane of erythrocytes infected with *Plasmodium falciparum*. *Journal of Protozoology,* 32**,** 424-429.

AITMAN, T. J., COOPER, L. D., NORSWORTHY, P. J., WAHID, F. N., GRAY, J. K., CURTIS, B. R., MCKEIGUE, P. M., KWIATKOWSKI, D., GREENWOOD, B. M., SNOW, R. W., HILL, A. V. & SCOTT, J. (2000) Malaria susceptibility and CD36 mutation. *Nature,* 405**,** 1015-1016.

AL-KHEDERY, B. & ALLRED, D. R. (2006) Antigenic variation in *Babesia bovis* occurs through segmental gene conversion of the *ves* multigene family, within a bidirectional locus of active transcription. *Molecular Microbiology,* 59**,** 402-414.

AL-KHEDERY, B., BARNWELL, J. W. & GALINSKI, M. R. (1999) Antigenic variation in malaria: a 3' genomic alteration associated with the expression of a *P. knowlesi* variant antigen. *Molecular Cell,* 3**,** 131-41.

ALARO, J. R., LYNCH, M. M. & BURNS, J. M. (2010) Protective immune responses elicited by immunization with a chimeric blood-stage malaria vaccine persist but are not boosted by *Plasmodium yoelii* challenge infection. *Vaccine,* 28**,** 6876-6884.

ALIBOLANDI, M. & MIRZAHOSEINI, H. (2011) Chemical Assistance in Refolding of Bacterial Inclusion Bodies. *Biochemistry Research International***,** e631607.

ALKAWA, M., MILLER, L. H. & RABBEGE, J. (1975) Caveola: vesicle complexes in the plasmalemma of erythrocytes infected by *Plasmodium vivax* and *P. cynomolgi*. Unique structures related to Schuffner's dots. *American Journal of Pathology,* 79**,** 385-300.

ALLEN, S. J., O'DONNELL, A., ALEXANDER, N. D. E., ALPERS, M. P., PETO, T. E. A., CLEGG, J. B. & WEATHERALL, D. J. (1997) α+-Thalassemia protects children against disease caused by other infections as well as malaria. *Proceedings of the National Academy of Sciences of the United States of America,* 94**,** 14736-14741.

ALLEN, S. J., O'DONNELL, A., ALEXANDER, N. D. E., MGONE, C. S., PETO, T. E. A., CLEGG, J. B., ALPERS, M. P. & WEATHERALL, D. J. (1999) Prevention of cerebral malaria in children in Papua New Guinea by Southeast Asian ovalocytosis band 3. *American Journal of Tropical Medicine and Hygiene,* 60**,** 1056-1060.

ALLRED, D. R., CARLTON, J. M. R., SATCHER, R. L., LONG, J. A., BROWN, W. C., PATTERSON, P. E., O'CONNOR, R. M. & STROUP, S. E. (2000) The *ves* multigene family of *B. bovis* encodes components of rapid antigenic variation at the infected erythrocyte surface. *Molecular Cell,* 5**,** 153-162.

ALLSOPP, C. E., SANNI, L. A., REUBSAET, L., NDUNGU, F., NEWBOLD, C., MWANGI, T., MARSH, K. & LANGHORNE, J. (2002) CD4 T cell responses to a variant antigen of the malaria parasite *Plasmodium falciparum*, erythrocyte membrane protein-1, in individuals living in malaria-endemic areas. *Journal of Infectious Diseases,* 185**,** 812-9.

ALONSO, P. L. (2006) Malaria: Deploying a candidate vaccine (RTS,S/AS02A) for an old scourge of humankind. *International Microbiology,* 9**,** 83-93.

ALTSCHUL, S. F., GISH, W., MILLER, W., MYERS, E. W. & LIPMAN, D. J. (1990) Basic local alignment search tool. *Journal of Molecular Biology,* 215**,** 403-410.

AMANTE, F. H. & GOOD, M. F. (1997) Prolonged Th1-like response generated by a *Plasmodium yoelii*-specific T cell clone allows complete clearance of infection in reconstituted mice. *Parasite Immunology,* 19**,** 111-126.

ANDERS, R. F., CREWTHER, P. E., EDWARDS, S., MARGETTS, M., MATTHEW, M. L., POLLOCK, B. & PYE, D. (1998) Immunisation with recombinant AMA-1 protects mice against infection with *Plasmodium chabaudi*. *Vaccine,* 16**,** 240-7.

ANDERSON, T., J., C., SU, X.-Z., BOCKAIRE, M., LAGOG, M. & DAY, K., P. (1999) Twelve microsatellite markers for characterisation of *Plasmodium falciparum* from finger prick blood samples. . *Parasitology,* 119**,** 113-125.

ANISIMOVA, M. & GASCUEL, O. (2006) Approximate Likelihood-Ratio Test for Branches: A Fast, Accurate, and Powerful Alternative. *Systematic Biology.,* 55**,** 539-552.

ANSARI, F. A., KUMAR, N., SUBRAMANYAM, M. B., GNANAMANI, M. & RAMACHANDRAN, S. (2008) MAAP: Malarial adhesins and adhesin-like proteins predictor. *Proteins: Structure, Function and Genetics,* 70**,** 659-666.

ANSTEY, N. M. & PRICE, R. N. (2007) Improving case definitions for severe malaria. *PLoS Medicine,* 4**,** 1291-1292.

ANSTEY, N. M., WEINBERG, J. B., HASSANALI, M. Y., MWAIKAMBO, E. D., MANYENGA, D., MISUKONIS, M. A., ARNELLE, D. R., HOLLIS, D., MCDONALD, M. I. & GRANGER, D. L. (1996) Nitric oxide in Tanzanian children with malaria: Inverse relationship between malaria severity and nitric oxide production/nitric oxide synthase type 2 expression. *Journal of Experimental Medicine,* 184**,** 557-567.

ASAMI, M., OWHASHI, M., ABE, T. & NAWA, Y. (1992) A comparative study of the kinetic changes of hemopoietic stem cells in mice infected with lethal and non-lethal malaria. *International Journal for Parasitology,* 22**,** 43-47.

AVRIL, M., CARTWRIGHT, M. M., HATHAWAY, M. J., HOMMEL, M., ELLIOTT, S. R., WILLIAMSON, K., NARUM, D. L., DUFFY, P. E., FRIED, M., BEESON, J. G. & SMITH, J. D. (2010) Immunization with VAR2CSA-DBL5 recombinant protein elicits broadly cross-reactive antibodies to placental *Plasmodium falciparum*-infected erythrocytes. *Infection and Immunity,* 78**,** 2248-2256.

AWANDARE, G. A., GOKA, B., BOEUF, P., TETTEH, J. K., KURTZHALS, J. A., BEHR, C. & AKANMORI, B. D. (2006) Increased levels of inflammatory mediators in children with severe *Plasmodium falciparum* malaria with respiratory distress. *Journal of Infectious Diseases,* 194**,** 1438-1446.

BACHMANN, A., ESSER, C., PETTER, M., PREDEHL, S., VON KALCKREUTH, V., SCHMIEDEL, S., BRUCHHAUS, I. & TANNICH, E. (2009) Absence of erythrocyte sequestration and lack of multicopy gene family expression in *Plasmodium falciparum* from a splenectomized malaria patient. *PLoS One,* 4.

BAER, K., KLOTZ, C., KAPPE, S. H. I., SCHNIEDER, T. & FREVERT, U. (2007) Release of hepatic *Plasmodium yoelii* merozoites into the pulmonary microvasculature. *PLoS Pathogens,* 3**,** 1651-1668.

BAILEY, T. L. & ELKAN, C. (1994) Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology.* Menlo Park, California, AAAI Press.

BAIRD, J. K., KRISIN, BARCUS, M. J., ELYAZAR, I. R. F., BANGS, M. J., MAGUIRE, J. D., FRYAUFF, D. J., RICHIE, T. L., SEKARTUTI & KALALO, W. (2003) Onset of clinical immunity to *Plasmodium falciparum* among Javanese migrants to Indonesian Papua. *Annals of Tropical Medicine and Parasitology,* 97**,** 557-564.

BALAJI, S., MADAN BABU, M., IYER, L. M. & ARAVIND, L. (2005) Discovery of the principal specific transcription factors of Apicomplexa and their implication for the evolution of the AP2-integrase DNA binding domains. *Nucleic Acids Research,* 33**,** 3994-4006.

BARNWELL, J. W., HOWARD, R. J., COON, H. G. & MILLER, L. H. (1983) Splenic requirement for antigenic variation and expression of the variant antigen on the erythrocyte membrane in cloned *Plasmodium knowlesi* malaria. *Infection and Immunity,* 40**,** 985-94.

BARNWELL, J. W., HOWARD, R. J. & MILLER, L. H. (1982) Altered expression of *Plasmodium knowlesi* variant antigen on the erythrocyte membrane in splenectomized rhesus monkeys. *Journal of Immunology,* 128**,** 224-6.

BARRY, A. E., LELIWA-SYTEK, A., TAVUL, L., IMRIE, H., MIGOT-NABIAS, F., BROWN, S. M., MCVEAN, G. A. V. & DAY, K. P. (2007) Population genomics of the immune evasion (*var*) genes of *Plasmodium falciparum*. *PLoS Pathogens,* 3.

BARUCH, D. I., GORMELY, J. A., MA, C., HOWARD, R. J. & PASLOSKE, B. L. (1996) *Plasmodium falciparum* erythrocyte membrane protein 1 is a parasitized erythrocyte receptor for adherence to CD36, thrombospondin, and intercellular adhesion molecule 1. *Proceedings of the National Academy of Sciences of the United States of America,* 93**,** 3497-502.

BASTIN, P., BAGHERZADEH, A., MATTHEWS, K. R. & GULL, K. (1996) A novel epitope tag system to study protein targeting and organelle biogenesis in *Trypanosoma brucei*. *Molecular and Biochemical Parasitology,* 77**,** 235-239.

BEESON, J., G., OSIER, F., H. & ENGWERDA, C., R. (2008) Recent insights into humoral and cellular immune responses against malaria. *Trends in Parasitology,* 24**,** 578-84.

BELYAEV, N. N., BROWN, D. E., DIAZ, A. I. G., RAE, A., JARRA, W., THOMPSON, J., LANGHORNE, J. & POTOCNIK, A. J. (2010) Induction of an IL7-R+ c-Kit hi myelolymphoid progenitor critically dependent on IFN-γ signaling during acute malaria. *Nature Immunology,* 11**,** 477-485.

BERENDT, A. R., TURNER, G. D. H. & NEWBOLD, C. I. (1994) Cerebral malaria: The sequestration hypothesis. *Parasitology Today,* 10**,** 412-414.

BERNARDS, A., VAN DER PLOEG, L. H. T. & FRASCH, A. C. C. (1981) Activation of trypanosome surface glycoprotein genes involves a duplication-transposition leading to an altered 3' end. *Cell,* 27**,** 497-505.

BHATTACHARJEE, S., HILLER, N. L., LIOLIOS, K., WIN, J., KANNEGANTI, T. D., YOUNG, C., KAMOUN, S. & HALDAR, K. (2006) The malarial host-targeting signal is conserved in the Irish potato famine pathogen. *PLoS Pathogens,* 2**,** 453-465.

BIGGS, B. A., ANDERS, R. F., DILLON, H. E., DAVERN, K. M., MARTIN, M., PETERSEN, C. & BROWN, G. V. (1992) Adherence of infected erythrocytes to venular endothelium selects for antigenic variants of *Plasmodium falciparum*. *Journal of Immunology,* 149**,** 2047-54.

BIGGS, B. A., GOOZE, L., WYCHERLEY, K., WOLLISH, W., SOUTHWELL, B., LEECH, J. H. & BROWN, G. V. (1991) Antigenic variation in *Plasmodium falciparum*. *Proceedings of the National Academy of Sciences of the United States of America,* 88**,** 9171-9174.

BLACKMAN, M. J., HEIDRICH, H. G., DONACHIE, S., JANA, S., MCBRIDE, S. & HOLDER, A. A. (1990) A Single Fragment of a Malaria Merozoite Surface Protein Remains on the Parasite During Red Cell Invasion and Is the Target of Invasion-inhibiting Antibodies. *Journal of Exprimental Medicine,* 172 379 - 382.

BLATTNER, F. R. & TUCKER, P. W. (1984) The molecular biology of immunoglobulin D. *Nature,* 307**,** 417-422.

BLISNICK, T., MORALES BETOULLE, M. E., BARALE, J. C., UZUREAU, P., BERRY, L., DESROSES, S., FUJIOKA, H., MATTEI, D. & BRAUN BRETON, C. (2000) *Pf*SBP1, a Maurer's cleft *Plasmodium falciparum* protein, is associated with the erythrocyte skeleton. *Molecular and Biochemical Parasitology,* 111**,** 107-121.

BLYTHE, J. E., XUE, Y. Y., KUSS, C., BOZDECH, Z., HOLDER, A. A., MARSH, K., LANGHORNE, J. & PREISER, P. R. (2008) *Plasmodium falciparum* STEVOR proteins are highly expressed in patient isolates and located in the surface membranes of infected red blood cells and the apical tips of merozoites. *Infection and Immunity,* 76**,** 3329-3336.

BODDEY, J. A., MORITZ, R. L., SIMPSON, R. J. & COWMAN, A. F. (2009) Role of the *Plasmodium* export element in trafficking parasite proteins to the infected erythrocyte. *Traffic,* 10**,** 285-299.

BOOKOUT, A. L., CUMMINS, C. L., MANGELSDORF, D. J., PESOLA, J. M. & KRAMER, M. F. (2006) High-throughput real-time quantitative reverse transcription PCR. IN AUSUBEL, F. M. (Ed.) *Current protocols in molecular biology* New York Wiley.

BOYLE, D. B., NEWBOLD, C. I., SMITH, C. C. & BROWN, K. N. (1982) Monoclonal antibodies that protect *in vivo* against *Plasmodium chabaudi* recognize a 250,000-dalton parasite polypeptide. *Infection and Immunity,* 38**,** 94-102.

BOZDECH, Z., LLINAS, M., PULLIAM, B. L., WONG, E. D., ZHU, J. & DERISI, J. L. (2003a) The transcriptome of the intraerythrocytic developmental cycle of *Plasmodium falciparum*. *PLoS Biology,* 1.

BOZDECH, Z., MOK, S., HU, G., IMWONG, M., JAIDEE, A., RUSSELL, B., GINSBURG, H., NOSTEN, F., DAY, N. P., WHITE, N. J., CARLTON, J. M. & PREISER, P. R. (2008) The transcriptome of *Plasmodium vivax* reveals divergence and diversity of transcriptional regulation in malaria parasites. *Proceedings of the National Academy of Sciences of the United States of America,* 105**,** 16290 - 16295.

BOZDECH, Z., ZHU, J., JOACHIMIAK, M. P., COHEN, F. E., PULLIAM, B. & DERISI, J. L. (2003b) Expression profiling of the schizont and trophozoite stages of *Plasmodium falciparum* with a long-oligonucleotide microarray. *Genome Biology,* 4**,** R9.

BRADY, C. P., SHIMP, R. L., MILES, A. P., WHITMORE, M. & STOWERS, A. W. (2001) High-level production and purification of P30P2MSP1(19), an important vaccine antigen for malaria, expressed in the methylotropic yeast *Pichia pastoris*. *Protein Expression and Purification,* 23**,** 468-475.

BRANCH, O., UDHAYAKUMAR, V., HIGHTOWER, A., OLOO, A., HAWLEY, W., NAHLEN, B., BLOLAND, P., KASLOW, D. & LAL, A. (1998) A longitudinal investigation of IgG and IgM antibody responses to the merozoite surface protein-1 19-kilodalton domain of *Plasmodium falciparum* in pregnant women and infants: associations with febrile illness, parasitemia and anemia. *American Journal of Tropical Medicine and Hygiene,* 58**,** 211-219.

BRAND, S. (2009) Recent advances in basic science Crohn's disease: Th1, Th17 or both? The change of a paradigm: new immunological and genetic insights implicate Th17 cells in the pathogenesis of Crohn's disease. *Gut,* 58**,** 1152-1167.

BRANNAN, L. R., TURNER, C. M. R. & PHILLIPS, R. S. (1994) Malaria parasites undergo antigenic variation at high rates in vivo. *Proceedings of the Royal Society B: Biological Sciences,* 256**,** 71-75.

BRIDGES, D. J., BUNN, J., VAN MOURIK, J. A., GRAU, G., PRESTON, R. J. S., MOLYNEUX, M., COMBES, V., O'DONNELL, J. S., DE LAAT, B. & CRAIG, A. (2010) Rapid activation of endothelial cells enables *Plasmodium falciparum* adhesion to platelet-decorated von Willebrand factor strings. *Blood,* 115**,** 1472-1474.

BRIERLEY, R. A., DAVIS, G. R. & HOLTZ, G. C. (1994) Production of Insulin-Like Growth Factor-1 in Methylotrophic Yeast Cells. *United States Patent.*

BROOKMAN, J. L., STOTT, A. J., CHEESEMAN, P. J., BURNS, N. R., ADAMS, S. E., KINGSMAN, A. J. & GULL, K. (1995) An immunological analysis of Ty1 virus-like particle structure. *Virology,* 207**,** 59-67.

BROWN, H., TURNER, G., ROGERSON, S., TEMBO, M., MWENECHANYA, J., MOLYNEUX, M. & TAYLOR, T. (1999) Cytokine expression in the brain in human cerebral malaria. *Journal of Infectious Diseases,* 180**,** 1742-1746.

BROWN, K., N. (1976) Resistance to malaria. IN COHEN, S. & SADUN, E. H. (Eds.) *Immunity to parasitic diseases.* Oxford, Blackwell Scientific.

BROWN, K., N. & BROWN, I., N. (1965) Immunity to malaria: antigenic variation in chronic infections of *Plasmodium knowlesi*. *Nature,* 208**,** 1286-1288.

BRYANT, D. & MOULTON, V. (2004) Neighbor-net: an agglomerative method for the construction of phylogenetic networks. *Molecular Biology and Evolution,* 21**,** 255–65.

BUFFET, P., A. , SAFEUKUI, I., DEPLAINE, G., BROUSSE, V., PRENDKI, V., THELLIER, M., TURNER, G., D. & MERCEREAU-PUIJALON, O. (2011) The pathogenesis of *Plasmodium falciparum* malaria in humans: insights from splenic physiology. *Blood,* 117**,** 381-392.

BULL, P., C. & MARSH, K. (2002) The role of antibodies to *Plasmodium falciparum*-infected-erythrocyte surface antigens in naturally acquired immunity to malaria. *Trends in Microbiology,* 10**,** 55-58.

BULL, P. C., BERRIMAN, M., KYES, S., QUAIL, M. A., HALL, N., KORTOK, M. M., MARSH, K. & NEWBOLD, C. I. (2005a) *Plasmodium falciparum* variant surface antigen expression patterns during malaria. *PLoS Pathogens,* 1**,** e26.

BULL, P. C., BUCKEE, C. O., KYES, S., KORTOK, M. M., THATHY, V., GUYAH, B., STOUTE, J. A., NEWBOLD, C. I. & MARSH, K. (2008) *Plasmodium falciparum* antigenic variation. Mapping mosaic *var* gene sequences onto a network of shared, highly polymorphic sequence blocks. *Molecular Microbiology,* 68**,** 1519-1534.

BULL, P. C., KORTOK, M., KAI, O., NDUNGU, F., ROSS, A., LOWE, B. S., NEWBOLD, C. I. & MARSH, K. (2000) *Plasmodium falciparum*-infected erythrocytes: agglutination by diverse Kenyan plasma is associated with severe disease and young host age. *Journal of Infectious Diseases,* 182**,** 252-9.

BULL, P. C., KYES, S., BUCKEE, C. O., MONTGOMERY, J., KORTOK, M. M., NEWBOLD, C. I. & MARSH, K. (2007) An approach to classifying sequence tags sampled from *Plasmodium falciparum var* genes. *Molecular and Biochemical Parasitology,* 154**,** 98-102.

BULL, P. C., LOWE, B. S., KORTOK, M., MOLYNEUX, C. S., NEWBOLD, C. I. & MARSH, K. (1998) Parasite antigens on the infected red cell surface are targets for naturally acquired immunity to malaria. *Nature Medicine,* 4**,** 358-60.

BULL, P. C., PAIN, A., NDUNGU, F. M., KINYANJUI, S. M., ROBERTS, D. J., NEWBOLD, C. I. & MARSH, K. (2005b) *Plasmodium falciparum* antigenic variation: relationships between *in vivo* selection, acquired antibody response, and disease severity. *Journal of Infectious Diseases,* 192**,** 1119-26.

BULTRINI, E., BRICK, K., MUKHERJEE, S., ZHANG, Y., SILVESTRINI, F., ALANO, P. & PIZZI, E. (2009) Revisiting the *Plasmodium falciparum* RIFIN family: From comparative genomics to 3D-model prediction. *BMC Genomics,* 10**,** 445.

BURNS, J., M. JR. , FLAHERTY, P. R., ROMERO, M. M. & WEIDANZ, W. P. (2003) Immunization against *Plasmodium chabaudi* malaria using combined formulations of apical membrane antigen-1 and merozoite surface protein-1. *Vaccine,* 21**,** 1843-52.

BURNS, J. M. J., FLAHERTY, P. R., NANAVATI, P. & WEIDANZ, W. P. (2004) Protection against *Plasmodium chabaudi* malaria induced by immunization with apical membrane antigen 1 and merozoite surface protein 1 in the absence of gamma interferon or interleukin-4. *Infection and Immunity,* 72**,** 5605-5612.

BUSTIN, S. A., BENES, V., GARSON, J. A., HELLEMANS, J., HUGGETT, J., KUBISTA, M., MUELLER, R., NOLAN, T., PFAFFL, M. W., SHIPLEY, G. L., VANDESOMPELE, J. & WITTWER, C. T. (2009) The MIQE guidelines: Minimum

information for publication of quantitative real-time PCR experiments. *Clinical Chemistry,* 55**,** 611-622.

CABRAL, F. J. & WUNDERLICH, G. (2009) Transcriptional memory and switching in the *Plasmodium falciparum rif* gene family. *Molecular and Biochemical Parasitology,* 168**,** 186-90.

CADMAN, E. T., ABDALLAH, A. Y., VOISINE, C., SPONAAS, A. M., CORRAN, P., LAMB, T., BROWN, D., NDUNGU, F. & LANGHORNE, J. (2008) Alterations of splenic architecture in malaria are induced independently of toll-like receptors 2, 4, and 9 or MyD88 and may affect antibody affinity. *Infection and Immunity,* 76**,** 3924-3931.

CARLSON, J., HELMBY, H., HILL, A. V. S., BREWSTER, D., GREENWOOD, B. M. & WAHLGREN, M. (1990) Human cerebral malaria: Association with erythrocyte rosetting and lack of anti-rosetting antibodies. *Lancet,* 336**,** 1457-1460.

CARLTON, J. M., ADAMS, J. H., SILVA, J. C., BIDWELL, S. L., LORENZI, H., CALER, E., CRABTREE, J., ANGIUOLI, S. V., MERINO, E. F., AMEDEO, P., CHENG, Q., COULSON, R. M. R., CRABB, B. S., DEL PORTILLO, H. A., ESSIEN, K., FELDBLYUM, T. V., FERNANDEZ-BECERRA, C., GILSON, P. R., GUEYE, A. H., GUO, X., KANG'A, S., KOOIJ, T. W. A., KORSINCZKY, M., MEYER, E. V. S., NENE, V., PAULSEN, I., WHITE, O., RALPH, S. A., REN, Q., SARGEANT, T. J., SALZBERG, S. L., STOECKERT, C. J., SULLIVAN, S. A., YAMAMOTO, M. M., HOFFMAN, S. L., WORTMAN, J. R., GARDNER, M. J., GALINSKI, M. R., BARNWELL, J. W. & FRASER-LIGGETT, C. M. (2008) Comparative genomics of the neglected human malaria parasite *Plasmodium vivax*. *Nature,* 455**,** 757-763.

CARLTON, J. M., ANGIUOLI, S. V., SUH, B. B., KOOIJ, T. W., PERTEA, M., SILVA, J. C., ERMOLAEVA, M. D., ALLEN, J. E., SELENGUT, J. D., KOO, H. L., PETERSON, J. D., POP, M., KOSACK, D. S., SHUMWAY, M. F., BIDWELL, S. L., SHALLOM, S. J., VAN AKEN, S. E., RIEDMULLER, S. B., FELDBLYUM, T. V., CHO, J. K., QUACKENBUSH, J., SEDEGAH, M., SHOAIBI, A., CUMMINGS, L. M., FLORENS, L., YATES, J. R., RAINE, J. D., SINDEN, R. E., A., H. M., CUNNINGHAM, D. A., PREISER, P. R., BERGMAN, L. W., VAIDYA, A. B., H, V. L. L., JANSE, C. J., WATERS, A. P., SMITH, H. O., WHITE, O. R., SALZBERG, S. L., VENTER, J. C., FRASER, C. M., HOFFMAN, S. L., GARDNER, M. J. & CARUCCI, D. J. (2002) Genome sequence and comparative analysis of the model rodent malaria parasite *Plasmodium yoelii yoelii*. *Nature,* 419**,** 512-9.

CARRIO, M. M. & VILLAVERDE, A. (2002) Construction and deconstruction of bacterial inclusion bodies. *Journal of Biotechnology,* 96**,** 3-12.

CARVALHO, B. O., LOPES, S. C. P., NOGUEIRA, P. A., ORLANDI, P. P., BARGIERI, D. Y., BLANCO, Y. C., MAMONI, R., LEITE, J. A., RODRIGUES, M. M., SOARES, I. S., OLIVEIRA, T. R., WUNDERLICH, G., LACERDA, M. V. G., DEL PORTILLO, H. A., ARAÚJO, M. O. G., RUSSELL, B., SUWANARUSK, R., SNOUNOU, G., RÉNIA, L. & COSTA, F. T. M. (2010) On the cytoadhesion of *Plasmodium vivax*-infected erythrocytes. *Journal of Infectious Diseases,* 202**,** 638-647.

CARVER, T., BERRIMAN, M., TIVEY, A., PATEL, C., BÖHME, U., BARRELL, B., G. , PARKHILL, J. & RAJANDREAM, M. A. (2008) Artemis and ACT: Viewing, annotating and comparing sequences stored in a relational database. *Bioinformatics*.

CAVINATO, R. A., BASTOS, K. R., SARDINHA, L. R., ELIAS, R. M., ALVAREZ, J. M. & D'IMPÉRIO LIMA, M. R. (2001) Susceptibility of the different developmental stages of the asexual (schizogonic) erythrocyte cycle of *Plasmodium chabaudi chabaudi* to hyperimmune serum, immunoglobulin (Ig)G1, IgG2a and F(ab')2 fragments. *Parasite Immunology,* 23**,** 587-597.

CHAM, G. K., TURNER, L., KURTIS, J. D., MUTABINGWA, T., FRIED, M., JENSEN, A. T., LAVSTSEN, T., HVIID, L., DUFFY, P. E. & THEANDER, T. G. (2010) Hierarchical, domain type-specific acquisition of antibodies to *Plasmodium falciparum* erythrocyte membrane protein 1 in Tanzanian children. *Infection and Immunity,* 78**,** 4653-9.

CHAM, G. K. K., TURNER, L., LUSINGU, J., VESTERGAARD, L., MMBANDO, B. P., KURTIS, J. D., JENSEN, A. T. R., SALANTI, A., LAVSTSEN, T. & THEANDER, T. G. (2009) Sequential, Ordered Acquisition of Antibodies to *Plasmodium falciparum* Erythrocyte Membrane Protein 1 Domains. *Journal of Immunology,* 183**,** 3356-3363.

CHANG, H. H., FALICK, A. M., CARLTON, P. M., SEDAT, J. W., DERISI, J. L. & MARLETTA, M. A. (2008) N-terminal processing of proteins exported by malaria parasites. *Molecular and Biochemical Parasitology,* 160**,** 107-115.

CHANG, S. P., GIBSON, H. L., LEE-NG, C. T., BARR, P. J. & HUI, G. S. (1992) A carboxyl-terminal fragment of *Plasmodium falciparum* gp195 expressed by a recombinant baculovirus induces antibodies that completely inhibit parasite growth. *Journal of Immunology,* 149**,** 548-55.

CHATTOPADHYAY, R., SHARMA, A., SRIVASTAVA, V. K., PATI, S. S., SHARMA, S. K., DAS, B. S. & CHITNIS, C. E. (2003) *Plasmodium falciparum* infection elicits both variant-specific and cross-reactive antibodies against variant surface antigens. *Infection and Immunity,* 71**,** 597-604.

CHEN, Q. (2007) The naturally acquired immunity in severe malaria and its implication for a *Pf*EMP-1 based vaccine. *Microbes and Infection,* 9**,** 777-783.

CHEN, Q., FERNANDEZ, V., SUNDSTRÖM, A., SCHLICHTHERLE, M., DATTA, S., HAGBLOM, P. & WAHLGREN, M. (1998) Developmental selection of *var* gene expression in *Plasmodium falciparum*. *Nature,* 394**,** 392-395.

CHEN, Q., HEDDINI, A., BARRAGAN, A., FERNANDEZ, V., PEARCE, S. F. A. & WAHLGREN, M. (2000) The semiconserved head structure of *Plasmodium falciparum* erythrocyte membrane protein 1 mediates binding to multiple independent host receptors. *Journal of Experimental Medicine,* 192**,** 1-9.

CHENG, Q., CLOONAN, N., FISCHER, K., THOMPSON, J., WAINE, G., LANZER, M. & SAUL, A. (1998) *Stevor* and *rif* are *Plasmodium falciparum* multicopy gene families which potentially encode variant antigens. *Molecular and Biochemical Parasitology,* 97**,** 161-76.

CHOOKAJORN, T., PONSUWANNA, P. & CUI, L. (2008) Mutually exclusive *var* gene expression in the malaria parasite: multiple layers of regulation. *Trends in Parasitology,* 24**,** 455-461.

CHRISTOPHERS, S. R. & FULTON, J. D. (1939) Experiments with isolated malaria parasites (*Plasmodium knowlesi*) free from red cells. . *Annals of tropical medicine and parasitology,* 33**,** 161-170.

CLARE, J. J., RAYMENT, F. B., BALLANTINE, S. P., SREEKRISHNA, K. & ROMANOS, M. A. (1991a) High-level Expression of Tetanus Toxin Fragment c in *Pichia pastoris* Strains Containing Multiple Tandem Integrations of the Gene. *Biotechnology,* 9**,** 455-460.

CLARE, J. J., ROMANOS, M. A., RAYMENT, F. B., ROWEDDER, J. E., SMITH, M. A., PAYNE, M. M., SREEKRISHNA, K. & HENWOOD, C. A. (1991b) Production of Epidermal Growth Factor in Yeast: High-Level Secretion Using *Pichia pastoris* Strains Containing Multiple Gene Copies. *Gene,* 105**,** 202-212.

CLARK, I. A., AWBURN, M. M., WHITTEN, R. O., HARPER, C. G., LIOMBA, N. G., MOLYNEUX, M. E. & TAYLOR, T. E. (2003) Tissue distribution of migration inhibitory factor and inducible nitric oxide synthase in falciparum malaria and sepsis in African children. *Malaria journal [electronic resource],* 2**,** 6.

CLARK, I. A. & COWDEN, W. B. (2003) The pathophysiology of *P. falciparum* malaria. *Pharmacology and Therapeutics,* 99**,** 221-260.

COATES, J. C. (2003) Armadillo repeat proteins: Beyond the animal kingdom. *Trends in Cell Biology,* 13**,** 463-471.

COCKBURN, I. A., MACKINNON, M. J., O'DONNELL, A., ALLEN, S. J., MOULDS, J. M., BAISOR, M., BOCKARIE, M., REEDER, J. C. & ROWE, J. A. (2004) A human complement receptor 1 polymorphism that reduces *Plasmodium falciparum* rosetting confers protection against severe malaria. *Proceedings of the National Academy of Sciences of the United States of America,* 101**,** 272-277.

COHEN, S. (1979) Immunity to malaria. *Proceedings of the Royal Society - Biological Sciences (Series B),* 203**,** 323-345.

COHEN, S., MCGREGOR, I. A. & CARRINGTON, S. (1961) Gamma-globulin and acquired immunity to human malaria. *Nature,* 192**,** 733-737.

COLER, R. N., CARTER, D., FRIEDE, M. & REED, S. G. (2009) Adjuvants for malaria vaccines. *Parasite Immunology,* 31**,** 520-528.

COLLINS, J. T. & DUNNICK, W. A. (1993) Germline transcripts of the murine immunoglobulin gamma 2a gene: structure and induction by IFN-gamma. . *International Immunology,* 5**,** 885-91.

COMBES, V., EL-ASSAAD, F., FAILLE, D., JAMBOU, R., HUNT, N. H. & GRAU, G. E. R. (2010) Microvesiculation and cell interactions at the brain-endothelial interface in cerebral malaria pathogenesis. *Progress in Neurobiology,* 91**,** 140-151.

CONWAY, D. J., CAVANAGH, D. R., TANABE, K., ROPER, C., MIKES, Z. S., SAKIHAMA, N., BOJANG, K. A., ODUOLA, A. M. J., KREMSNER, P. G., ARNOT, D. E., GREENWOOD, B. M. & MCBRIDE, J. S. (2000) A principal target of human immunity to malaria identified by molecular population genetic and immunological analyses. *Nature Medicine,* 6.

COOPER, N. R. (1985) The classical complement pathway: Activation and regulation of the first complement component. *Advances in Immunology,* 37**,** 151-216.

CORRAN, P., H., O'DONNELL, R., A., TODD, J., UTHAIPIBULL, C., HOLDER, A. A., CRABB, B. S. & RILEY, E. M. (2004) The fine specificity, but not the invasion inhibitory activity, of 19-kilodalton merozoite surface protein 1-specific antibodies is associated with resistance to malarial parasitemia in a cross-sectional survey in The Gambia. . *Infection and Immunity,* 72**,** 6185-6189.

CORREDOR, V., MEYER, E. V. S., LAPP, S., CORREDOR-MEDINA, C., HUBER, C. S., EVANS, A. G., BARNWELL, J. W. & GALINSKI, M. R. (2004) A *SICAvar* switching event in *Plasmodium knowlesi* is associated with the DNA rearrangement of conserved 3' non-coding sequences. *Molecular and Biochemical Parasitology,* 138**,** 37-49.

COWMAN, A. F. & CRABB, B. S. (2006) Invasion of red blood cells by malaria parasites. *Cell,* 124**,** 755-766.

COX, J., SEMOFF, S. & HOMMEL, M. (1987) *Plasmodium chabaudi*: A rodent malaria model for *in-vivo* and *in-vitro* cytoadherence of malaria parasites in the absence of knobs. *Parasite Immunology,* 9**,** 543-561.

COX-SINGH, J., DAVIS, T. M. E., LEE, K. S., SHAMSUL, S. S. G., MATUSOP, A., RATNAM, S., RAHMAN, H. A., CONWAY, D. J. & SINGH, B. (2008) *Plasmodium knowlesi* malaria in humans is widely distributed and potentially life threatening. *Clinical Infectious Diseases,* 46**,** 165-171.

CREGG, J. M., MADDEN, K. R., BARRINGER, K. J., THILL, G. & STILLMAN, C. A. (1989) Functional Characterization of the Two Alcohol Oxidase Genes from the Yeast, *Pichia pastoris*. *Molecular Cell Biology,* 9**,** 1316-1323.

CREWTHER, P., E. , MATTHEW, M. L., FLEGG, R. H. & ANDERS, R. F. (1996) Protective immune responses to apical membrane antigen 1 of *Plasmodium chabaudi* involve recognition of strain-specific epitopes. *Infection and Immunity,* 64**,** 3310-3317.

CROOKS, G., E. , HON, G., CHANDONIA, J., M. & BRENNER, S., E. (2004) WebLogo: A sequence logo generator. *Genome Research,* 14**,** 1188-1190.

CROSNIER, C., WANAGURU, M., MCDADE, B., RAYNER, J. & WRIGHT, G. (2010) A library of functional recombinant *Plasmodium falciparum* merozoite surface proteins *Malaria Journal,* 9**,** O8.

CROSS, G. A. M. (1996) Antigenic variation in trypanosomes: Secrets surface slowly. *BioEssays,* 18**,** 283-291.

CUNNINGHAM, D., FONAGER, J., JARRA, W., CARRET, C., PREISER, P. & LANGHORNE, J. (2009) Rapid changes in transcription profiles of the *Plasmodium yoelii yir* multigene family in clonal populations: Lack of epigenetic memory? *PLoS One,* 4.

CUNNINGHAM, D., LAWTON, J., JARRA, W., PREISER, P. & LANGHORNE, J. (2010) The *pir* multigene family of *Plasmodium*: Antigenic variation and beyond. *Molecular and Biochemical Parasitology,* 170**,** 65-73.

CUNNINGHAM, D. A., JARRA, W., KOERNIG, S., FONAGER, J., FERNANDEZ-REYES, D., BLYTHE, J. E., WALLER, C., PREISER, P. R. & LANGHORNE, J. (2005) Host immunity modulates transcriptional changes in a multigene family (*yir*) of rodent malaria. *Molecular Microbiology,* 58**,** 636-647.

DAHLBÄCK, M., LAVSTSEN, T., SALANTI, A., HVIID, L., ARNOT, D. E., THEANDER, T. G. & NIELSEN, M. A. (2007) Changes in *var* gene mRNA levels during erythrocytic development in two phenotypically distinct *Plasmodium falciparum* parasites. *Malaria Journal,* 6.

DAHNEKE, B. E. (Ed.) (1983) *Measurement of Suspended Particles by Quasielastic Light Scattering,* Hoboken, New Jersey, USA, John Wiley & Sons, Inc.

DALY, T. & LONG, C. (1993) A recombinant 15-kilodalton carboxyl terminal fragment of *Plasmodium yoelii yoelii* 17XL merozoite surface protein-1 induces a protective response in mice. *Infection and Immunity,* 61**,** 2462-2467.

DAS, P., GREWAL, J. S., MAHAJAN, B. & CHAUHAN, V. S. (2007) Comparison of cellular and humoral responses to recombinant protein and synthetic peptides of exon 2 region of *Plasmodium falciparum* erythrocyte membrane protein1 (*Pf*EMP1) among malaria patients from an endemic region. *Parasitology International,* 56**,** 51-59.

DAY, N. P. J., HIEN, T. T., SCHOLLAARDT, T., LOC, P. P., VAN CHUONG, L., CHAU, T. T. H., MAI, N. T. H., PHU, N. H., SINH, D. X., WHITE, N. J. & HO, M. (1999) The prognostic and pathophysiologic role of pro- and antiinflammatory cytokines in severe malaria. *Journal of Infectious Diseases,* 180**,** 1288-1297.

DE KONING-WARD, T. F., GILSON, P. R., BODDEY, J. A., RUG, M., SMITH, B. J., PAPENFUSS, A. T., SANDERS, P. R., LUNDIE, R. J., MAIER, A. G., COWMAN, A. F. & CRABB, B. S. (2009) A newly discovered protein export machine in malaria parasites. *Nature,* 459**,** 945-949.

DE SILVA, E. K., GEHRKE, A. R., OLSZEWSKI, K., LEÓN, I., CHAHAL, J. S., BULYK, M. L. & LLINÁS, M. (2008) Specific DNA-binding by Apicomplexan AP2 transcription factors. *Proceedings of the National Academy of Sciences of the United States of America,* 105**,** 8393-8398.

DE SOUZA, J. B., HAFALLA, J. C., RILEY, E. M. & COUPER, K. N. (2009) Cerebral malaria: Why experimental murine models are required to understand the pathogenesis of disease. *Parasitology***,** 1-18.

DEITSCH, K. W., CALDERWOOD, M. S. & WELLEMS, T. E. (2001) Malaria: Cooperative silencing elements in *var* genes. *Nature,* 412**,** 875-876.

DEITSCH, K. W., LUKEHART, S. A. & STRINGER, J. R. (2009) Common strategies for antigenic variation by bacterial, fungal and protozoan pathogens. *Nature Reviews Microbiology,* 7**,** 493-503.

DEL PORTILLO, H. A., FERNANDEZ-BECERRA, C., BOWMAN, S., OLIVER, K., PREUSS, M., SANCHEZ, C. P., SCHNEIDER, N. K., VILLALOBOS, J. M., RAJANDREAM, M. A., HARRIS, D., DA SILVA, L. H. P., BARRELL, B. & LANZER, M. (2001) A superfamily of variant genes encoded in the subtelomeric region of *Plasmodium vivax*. *Nature,* 410**,** 839-842.

DEL PORTILLO, H. A., LANZER, M., RODRIGUEZ-MALAGA, S., ZAVALA, F. & FERNANDEZ-BECERRA, C. (2004) Variant genes and the spleen in *Plasmodium vivax* malaria. *International Journal for Parasitology,* 34**,** 1547-1554.

DELORON, P. & CHOUGNET, C. (1992) Is immunity to malaria really short-lived? *Parasitology Today,* 8**,** 375-378.

DEVINE, D. V. (1991) The regulation of complement on cell surfaces. *Transfusion medicine reviews,* 5**,** 123-131.

DEY, A. K. & SRIVASTAVA, I. K. (2011) Novel adjuvants and delivery systems for enhancing immune responses induced by immunogens. *Expert Review of Vaccines,* 10**,** 227-251.

DI GIROLAMO, F., RAGGI, C., BIRAGO, C., PIZZI, E., LALLE, M., PICCI, L., PACE, T., BACHI, A., DE JONG, J., JANSE, C. J., WATERS, A. P., SARGIACOMO, M. & PONZI, M. (2008) *Plasmodium* lipid rafts contain proteins implicated in vesicular trafficking and signalling as well as members of the PIR superfamily, potentially implicated in host immune system interactions. *Proteomics,* 8**,** 2500-2513.

DIGGS, C. L. & OLSER, A. G. (1969) Humoral immunity in rodent malaria. II. Inhibition of parasitaemia by serum antibody. *Journal of Immunology,* 102**,** 298-305.

DOBBIE, M., CRAWLEY, J., WARUIRU, C., MARSH, K. & SURTEES, R. (2000) Cerebrospinal fluid studies in children with cerebral malaria: An excitotoxic mechanism? *American Journal of Tropical Medicine and Hygiene,* 62**,** 284-290.

DODOO, D., STAALSOE, T., GIHA, H., KURTZHALS, J. A., AKANMORI, B. D., KORAM, K., DUNYO, S., NKRUMAH, F. K., HVIID, L. & THEANDER, T. G. (2001) Antibodies to variant antigens on the surfaces of infected erythrocytes are associated with protection from malaria in Ghanaian children. *Infection and Immunity,* 69**,** 3713-8.

DODOO, D., THEANDER, T. G., KURTZHALS, J. A., KORAM, K., RILEY, E., AKANMORI, B. D., NKRUMAH, F. K. & HVIID, L. (1999) Levels of antibody to conserved parts of *Plasmodium falciparum* merozoite surface protein 1 in Ghanaian children are not associated with protection from clinical malaria. *Infection and Immunity,* 67**,** 2131-2137.

DONDORP, A. M., INCE, C., CHARUNWATTHANA, P., HANSON, J., VAN KUIJEN, A., FAIZ, M. A., RAHMAN, M. R., HASAN, M., BIN YUNUS, E., GHOSE, A., RUANGVEERAYUT, R., LIMMATHUROTSAKUL, D., MATHURA, K., WHITE, N. J. & DAY, N. P. J. (2008) Direct *in vivo* assessment of microcirculatory dysfunction in severe *P. falciparum* malaria. *Journal of Infectious Diseases,* 197**,** 79-84.

DOOLAN, D. L., DOBAN, C. & BAIRD, J. K. (2009) Acquired Immunity to Malaria. *Clinical Microbiology Reviews,* 22**,** 13-36.

DURAISINGH, M. T., VOSS, T. S., MARTY, A. J., DUFFY, M. F., GOOD, R. T., THOMPSON, J. K., FREITAS JR, L. H., SCHERF, A., CRABB, B. S. & COWMAN, A. F. (2005) Heterochromatin silencing and locus repositioning linked to regulation of virulence genes in *Plasmodium falciparum*. *Cell,* 121**,** 13-24.

DWORZAK, M. N., FRITSCH, G., BUCHINGER, P., FLEISCHER, C., PRINTZ, D., ZELLNER, A., SCHÖLLHAMMER, A., STEINER, G., AMBROS, P. F. & GADNER, H. (1994) Flow cytometric assessment of human MIC2 expression in bone marrow, thymus, and peripheral blood. *Blood* 83**,** 415-425.

DZIKOWSKI, R. & DEITSCH, K. W. (2008) Active transcription is required for maintenance of epigenetic memory in the malaria parasite *Plasmodium falciparum*. *Journal of Molecular Biology,* 382**,** 288-297.

DZIKOWSKI, R. & DEITSCH, K. W. (2009) Genetics of antigenic variation in *Plasmodium falciparum*. *Current Genetics,* 55**,** 103-110.

DZIKOWSKI, R., FRANK, M. & DEITSCH, K. (2006) Mutually exclusive expression of virulence genes by malaria parasites is regulated independently of antigen production. *PLoS pathogens,* 2.

DZIKOWSKI, R., LI, F., AMULIC, B., EISBERG, A., FRANK, M., PATEL, S., WELLEMS, T. E. & DEITSCH, K. W. (2007) Mechanisms underlying mutually exclusive expression of virulence genes by malaria parasites. *EMBO Reports,* 8**,** 959-965.

EATON, M., D. (1938) The agglutination of *Plasmodium knowlesi* by immune serum. *Journal of Experimental Medicine,* 67**,** 857-870.

EDDY, S., R. (1996) Hidden Markov Models. *Current Opinion in Structural Biology,* 6**,** 361-365.

EDGAR, R. C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research,* 32**,** 1792-1797.

EDOZIEN, J. C., GILLES, H. M. & UDEOZO, I. O. K. (1962) Adult and cord blood gamma-globulin and immunity to malaria in Nigerians. *Lancet,* 2**,** 951-955.

EGAN, A. F., BURGHAUS, P., DRUILHE, P., HOLDER, A. A. & RILEY, E. M. (1999) Human antibodies to the 19kDa C-terminal fragment of *Plasmodium falciparum* merozoite surface protein 1 inhibit parasite growth *in vitro*. *Parasite Immunology,* 21**,** 133–139.

EGAN, A. F., MORRIS, J., BARNISH, G., ALLEN, S., GREENWOOD, B. M., KASLOW, D. C., HOLDER, A. A. & RILEY, E. M. (1996) Clinical immunity to *Plasmodium falciparum* malaria is associated with serum antibodies to the 19-kDa C-terminal fragment of the merozoite surface antigen, *Pf*MSP-1 *Journal of Infectious Diseases,* 173**,** 765-769.

EL HASSAN, A. M. A., SAEED, A. M., FANDREY, J. & JELKMANN, W. (1997) Decreased erythropoietin response in *Plasmodium falciparum* malaria- associated anaemia. *European Journal of Haematology,* 59**,** 299-304.

ELLIOTT, S. R., PAYNE, P. D., DUFFY, M. F., BYRNE, T. J., THAM, W. H., ROGERSON, S. J., BROWN, G. V. & EISEN, D. P. (2007) Antibody recognition of heterologous variant surface antigens after a single *Plasmodium falciparum* infection in previously naive adults. *American Journal of Tropical Medicine and Hygiene,* 76**,** 860-4.

ELLIS, S. B., BRUST, P. F., KOUTZ, P. J., WATERS, A. F., HARPOLD, M. M. & GINGERAS, T. R. (1985) Isolation of Alcohol Oxidase and Two other Methanol Regulatable Genes from the Yeast, *Pichia pastoris*. *Molecular Cell Biology,* 5**,** 1111-1121.

ELSE, K. J. & FINKELMAN, F. D. (1998) Intestinal nematode parasites, cytokines and effector mechanisms. . *International Journal for Parasitology,* 28**,** 1145-58.

ENGLISH, M., WARUIRU, C., AMUKOYE, E., MURPHY, S., CRAWLEY, J., MWANGI, I., PESHU, N. & MARSH, K. (1996) Deep breathing in children with severe malaria: Indicator of metabolic acidosis and poor outcome. *American Journal of Tropical Medicine and Hygiene,* 55**,** 521-524.

ENGSTLER, M., PFOHL, T., HERMINGHAUS, S., BOSHART, M., WIEGERTJES, G., HEDDERGOTT, N. & OVERATH, P. (2007) Hydrodynamic Flow-Mediated Protein Sorting on the Cell Surface of Trypanosomes. *Cell,* 131**,** 505-515.

EPP, C., LI, F., HOWITT, C. A., CHOOKAJORN, T. & DEITSCH, K. W. (2009) Chromatin associated sense and antisense noncoding RNAs are transcribed from the *var* gene family of virulence genes of the malaria parasite *Plasmodium falciparum*. *RNA,* 15**,** 116-127.

FAIRLIE-CLARKE, K. J., LAMB, T. J., LANGHORNE, J., GRAHAM, A. L. & ALLEN, J. E. (2010) Antibody isotype analysis of malaria-nematode co-infection: Problems and solutions associated with cross-reactivity. *BMC Immunology,* 11.

FEACHEM, R. G. A., PHILLIPS, A. A., HWANG, J., COTTER, C., WIELGOSZ, B., GREENWOOD, B. M., SABOT, O., RODRIGUEZ, M. H., ABEYASINGHE, R. R., GHEBREYESUS, T. A. & SNOW, R. W. (2010) Shrinking the malaria map: progress and prospects. *Lancet,* 376**,** 1566-1578.

FELSENSTEIN, J. (1985) Confidence limits on phylogenies: An approach using the bootstrap. . *Evolution,* 39**,** 783-791.

FERNANDEZ, V., HOMMEL, M., CHEN, Q., HAGBLOM, P. & WAHLGREN, M. (1999) Small, clonally variant antigens expressed on the surface of the *Plasmodium falciparum* -infected erythrocyte are encoded by the *rif* gene family and are the target of human immune responses. *Journal of Experimental Medicine,* 190**,** 1393-404.

FERNANDEZ-BECERRA, C., PEIN, O., DE OLIVEIRA, T. R., YAMAMOTO, M. M., CASSOLA, A. C., ROCHA, C., SOARES, I. S., DE BRAGANÇA PEREIRA, C. A. & DEL PORTILLO, H. A. (2005) Variant proteins of *Plasmodium vivax* are not clonally expressed in natural infections. *Molecular Microbiology,* 58**,** 648-658.

FIELD, J. W. & REID, J. A. (1956) Malaria control in Malaya; an appreciation of the work of Sir Malcolm Watson. *Journal of Tropical Medicine and Hygiene,* 59**,** 23-27.

FINN, R. D., MISTRY, J., TATE, J., COGGILL, P., HEGER, A., POLLINGTON, J. E., GAVIN, O. L., GUNASEKARAN, P., CERIC, G., FORSLUND, K., HOLM, L.,

SONNHAMMER, E. L., EDDY, S. R. & BATEMAN, A. (2010) The Pfam protein families database. *Nucleic Acids Research,* 38**,** D211-22.

FISCHER, K., CHAVCHICH, M., HUESTIS, R., WILSON, D. W., KEMP, D. J. & SAUL, A. (2003) Ten families of variant genes encoded in subtelomeric regions of multiple chromosomes of *Plasmodium chabaudi*, a malaria species that undergoes antigenic variation in the laboratory mouse. *Molecular Microbiology,* 48**,** 1209-1223.

FLECK, J. D., KAUFFMANN, C., SPILKI, F., LENCINA, C. L., ROEHE, P. M. & GOSMANNA, G. (2006) Adjuvant activity of Quillaja brasiliensis saponins on the immune responses to bovine herpes virus type 1 in mice. *Vaccine,* 24**,** 7129-7134.

FLORENS, L., LIU, X., WANG, Y., YANG, S., SCHWARTZ, O., PEGLAR, M., CARUCCI, D. J., YATES III, J. R. & WU, Y. (2004) Proteomics approach reveals novel proteins on the surface of malaria-infected erythrocytes. *Molecular and Biochemical Parasitology,* 135**,** 1-11.

FLORENS, L., WASHBURN, M. P., RAINE, J. D., ANTHONY, R. M., GRAINGER, M., HAYNES, J. D., MOCH, J. K., MUSTER, N., SACCI, J. B., TABB, D. L., WITNEY, A. A., WOLTERS, D., WU, Y., GARDNER, M. J., HOLDER, A. A., SINDEN, R. E., YATES, J. R. & CARUCCI, D. J. (2002) A proteomic view of the *Plasmodium falciparum* life cycle. *Nature,* 419**,** 520-526.

FLUECK, C., BARTFAI, R., NIEDERWIESER, I., WITMER, K., ALAKO, B. T. F., MOES, S., BOZDECH, Z., JENOE, P., STUNNENBERG, H. G. & VOSS, T. S. (2010) A major role for the *Plasmodium falciparum* ApiAP2 protein PfSIP2 in chromosome end biology. *PLoS Pathogens,* 6.

FONAGER, J., CUNNINGHAM, D., JARRA, W., KOERNIG, S., HENNEMAN, A. A., LANGHORNE, J. & PREISER, P. (2007) Transcription and alternative splicing in the *yir* multigene family of the malaria parasite *Plasmodium y. yoelii*: Identification of motifs suggesting epigenetic and post-transcriptional control of RNA expression. *Molecular and Biochemical Parasitology,* 156**,** 1-11.

FRANK, M., DZIKOWSKI, R., AMULIC, B. & DEITSCH, K. (2007) Variable switching rates of malaria virulence genes are associated with chromosomal position. *Molecular Microbiology,* 64**,** 1486-1498.

FRANK, M., DZIKOWSKI, R., COSTANTINI, D., AMULIC, B., BERDOUGO, E. & DEITSCH, K. (2006) Strict pairing of *var* promoters and introns is required for *var* gene silencing in the malaria parasite *Plasmodium falciparum*. *Journal of Biological Chemistry,* 281**,** 9942-9952.

FRANK, M., KIRKMAN, L., COSTANTINI, D., SANYAL, S., LAVAZEC, C., TEMPLETON, T. J. & DEITSCH, K. W. (2008) Frequent recombination events generate diversity within the multi-copy variant antigen gene families of *Plasmodium falciparum*. *International Journal for Parasitology,* 38**,** 1099-109.

FRANKE-FAYARD, B., JANSE, C. J., CUNHA-RODRIGUES, M., RAMESAR, J., BÜSCHER, P., QUE, I., LÖWIK, C., VOSHOL, P. J., DEN BOER, M. A. M., VAN DUINEN, S. G., FEBBRAIO, M., MOTA, M. M. & WATERS, A. P. (2005) Murine malaria parasite sequestration: CD36 is the major receptor, but cerebral pathology is unlinked to sequestration. *Proceedings of the National Academy of Sciences of the United States of America,* 102**,** 11468-11473.

FRANKE-FAYARD, B., WATERS, A. P. & JANSE, C. J. (2006) Real-time *in vivo* imaging of transgenic bioluminescent blood stages of rodent malaria parasites in mice. *Nature Protocols,* 1**,** 476-485.

FREEMAN, R. R., TREJDOSIEWICZ, A. J. & CROSS, G. A. M. (1980) Protective monoclonal antibodies recognizing stage-specific merozoite antigens of a rodent malaria. *Nature,* 284**,** 366-8.

FREITAS-JUNIOR, L. H., BOTTIUS, E., PIRRIT, L. A., DEITSCH, K. W., SCHEIDIG, C., GUINET, F., NEHRBASS, U., WELLEMS, T. E. & SCHERF, A. (2000) Frequent ectopic recombination of virulence factor genes in telomeric chromosome clusters of *P. falciparum*. *Nature,* 407**,** 1018-1022.

FREMOUNT, H. N. & MILLER, L. H. (1975) Deep vascular schizogony in *Plasmodium fragile*: organ distribution and ultrastructure of erythrocytes adherent to vascular endothelium. *American Journal of Tropical Medicine and Hygiene,* 24**,** 1-8.

FREVERT, U., ENGELMANN, S., ZOUGBÉDÉ, S., STANGE, J., NG, B., MATUSCHEWSKI, K., LIEBES, L. & YEE, H. (2005) Intravital observation of *Plasmodium berghei* sporozoite infection of the liver. *PLoS Biology,* 3**,** 1034-1046.

FRIED, M. & DUFFY, P. E. (1996) Adherence of *Plasmodium falciparum* to chondroitin sulfate A in the human placenta. *Science,* 272**,** 1502-4.

GALLUP, J. L. & SACHS, J. D. (2001) The economic burden of malaria. *American Journal of Tropical Medicine and Hygiene,* 64**,** 85-96.

GAMAIN, B., MILLER, L. H. & BARUCH, D. I. (2001) The surface variant antigens of *Plasmodium falciparum* contain cross-reactive epitopes. *Proceedings of the National Academy of Sciences of the United States of America,* 98**,** 2664-9.

GAMBETTE, P. & HUSON, D. H. (2008) Improved layout of phylogenetic networks. *IEEE/ACM Trans Comput Biol Bioinform,* 5**,** 472-9.

GARCIA, J. E., PUENTES, A., CURTIDOR, H., VERA, R., RODRIGUEZ, L., VALBUENA, J., LOPEZ, R., OCAMPO, M., CORTES, J., VANEGAS, M., ROSAS, J., REYES, C. & PATARROYO, M. E. (2005) Peptides from the *Plasmodium falciparum* STEVOR putative protein bind with high affinity to normal human red blood cells. *Peptides,* 26**,** 1133-1143.

GARDNER, M. J., HALL, N., FUNG, E., WHITE, O., BERRIMAN, M., HYMAN, R. W., CARLTON, J. M., PAIN, A., NELSON, K. E., BOWMAN, S., PAULSEN, I. T., JAMES, K., EISEN, J. A., RUTHERFORD, K., SALZBERG, S. L., CRAIG, A., KYES, S., CHAN, M.-S., NENE, V., SHALLOM, S. J., SUH, B., PETERSON, J., ANGIUOLI, S., PERTEA, M., ALLEN, J., SELENGUT, J., HAFT, D., MATHER, M. W., VAIDYA, A. B., MARTIN, D. M. A., FAIRLAMB, A. H., FRAUNHOLZ, M. J., ROOS, D. S., RALPH, S. A., MCFADDEN, G. I., CUMMINGS, L. M., SUBRAMANIAN, M. G., MUNGALL, C., VENTER, J. C., CARUCCI, D. J., HOFFMAN, S. L., NEWBOLD, C., DAVIS, R. W., FRASER, C. M. & BARRELL, B. (2002) Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature,* 419**,** 498-511.

GARDNER, M. J., TETTELIN, H., CARUCCI, D. J., CUMMINGS, L. M., ARAVIND, L., KOONIN, E. V., SHALLOM, S., MASON, T., YU, K. & FUJII, C. (1998) Chromosome 2 sequence of the human malaria parasite *Plasmodium falciparum*. *Science,* 282**,** 1126-1132.

GASCUEL, O. (1997) BIONJ: an improved version of the NJ algorithm based on a simple model of sequence data. *Molecular Biology and Evolution,* 14**,** 685-695.

GASTEIGER, E., HOOGLAND, C., GATTIKER, A., DUVAUD, S., WILKINS, M. R., APPEL, R. D. & BAIROCH, A. (2005) Protein Identification and Analysis Tools on the ExPASy Server. IN WALKER, J. M. (Ed.) *The Proteomics Protocols Handbook.* Humana Press

GEHDE, N., HINRICHS, C., MONTILLA, I., CHARPIAN, S., LINGELBACH, K. & PRZYBORSKI, J. M. (2009) Protein unfolding is an essential requirement for transport across the parasitophorous vacuolar membrane of *Plasmodium falciparum*. *Molecular Microbiology,* 71**,** 613-628.

GHOSH, S., MALHOTRA, P., LALITHA, P. V., GUHA-MUKHERJEE, S. & CHAUHAN, V. S. (2002) Expression of *Plasmodium falciparum* C-terminal region of merozoite surface protein (*Pf*MSP119), a potential malaria vaccine candidate in tobacco. *Plant Science,* 162**,** 335-343.

GIHA, H. A., STAALSOE, T., DODOO, D., ELHASSAN, I. M., ROPER, C., SATTI, G. M., ARNOT, D. E., HVIID, L. & THEANDER, T. G. (1999) Overlapping antigenic repertoires of variant antigens expressed on the surface of erythrocytes infected by *Plasmodium falciparum*. *Parasitology,* 119 (Pt 1)**,** 7-17.

GILKS, C. F., WALLIKER, D. & NEWBOLD, C. I. (1990) Relationships between sequestration, antigenic variation and chronic parasitism in *Plasmodium chabaudi chabaudi* - a rodent malaria model. *Parasite Immunology,* 12**,** 45-64.

GILLES, H. M., FLETCHER, K. A., HENDRICKSE, R. G., LINDNER, R., REDDY, S. & ALLAN, N. (1967) Glucose-6-phosphate-dehydrogenase deficiency, sickling, and malaria in African children in South Western Nigeria. *Lancet,* 1**,** 138-140.

GOOD, M. F., STANISIC, D., XU, H., ELLIOTT, S. & WYKES, M. (2004) The immunological challenge to developing a vaccine to the blood stages of malaria parasites. *Immunological Reviews,* 201**,** 254-267.

GOWDA, D. C. & DAVIDSON, E. A. (1999) Protein glycosylation in the malaria parasite. *Parasitology Today,* 15**,** 147-152.

GOZALO, A., LUCAS, C., CACHAY, M., WELLDE, B. T., HALL, T., BELL, B., WOOD, J., WATTS, D., WOOSTER, M., LYON, J. A., MOCH, J. K., HAYNES, J. D., WILLIAMS, J. S., HOLLAND, C., WATSON, E., KESTER, K. E., KASLOW, D. C. & BALLOU, W. R. (1998) Passive transfer of growth-inhibitory antibodies raised against yeast-expressed recombinant *Plasmodium falciparum* merozoite surface protein-1(19). *American Journal of Tropical Medicine and Hygiene* 59**,** 991-7.

GRAU, G. E., FAJARDO, L. F. & PIGUET, P. F. (1987) Tumor necrosis factor (cachectin) as an essential mediator in murine cerebral malaria. *Science,* 237**,** 1210-1212.

GRAU, G. E., MACKENZIE, C. D., CARR, R. A., REDARD, M., PIZZOLATO, G., ALLASIA, C., CATALDO, C., TAYLOR, T. E. & MOLYNEUX, M. E. (2003) Platelet accumulation in brain microvessels in fatal pediatric cerebral malaria. *Journal of Infectious Diseases,* 187**,** 461-466.

GRAU, G. E., TAYLOR, T. E., MOLYNEUX, M. E., WIRIMA, J. J., VASSALLI, P., HOMMEL, M. & LAMBERT, P. H. (1989) Tumor necrosis factor and disease severity in children with *P. falciparum* malaria. *New England Journal of Medicine,* 320**,** 1586-1591.

GROUX, H. & GYSIN, J. (1990) Opsonisation as an effector mechanism in human protection against asexual blood stages of *Plasmodium falciparum*: functional role of IgG subclasses. *Research in Immunology,* 141**,** 532–542.

GRUN, J. L. & WEIDANZ, W. P. (1983) Antibody-independent immunity to re-infection malaria in B-cell deficient mice. *Infection and Immunity,* 41**,** 1197-1204.

GUERRA, C. A., HOWES, R. E., PATIL, A. P., GETHING, P. W., VAN BOECKEL, T. P., TEMPERLEY, W. H., KABARIA, C. W., TATEM, A. J., MANH, B. H., ELYAZAR, I. R. F., BAIRD, J. K., SNOW, R. W. & HAY, S. I. (2010) The international limits and population at risk of *Plasmodium vivax* transmission in 2009. *PLoS Neglected Tropical Diseases,* 4.

GUINDON, S. & GASCUEL, O. (2003) PhyML - A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology,* 52**,** 696-704.

GUINDON, S., LETHIEC, F., DUROUX, P. & GASCUEL, O. (2005) PHYML Online--a web server for fast maximum likelihood-based phylogenetic inference *Nucleic Acids Research,* 33 (Web Server issue)**,** W557-9. .

GUPTA, S., SNOW, R. W., DONNELLY, C. A., MARSH, K. & NEWBOLD, C. (1999) Immunity to non-cerebral severe malaria is acquired after one or two infections. *Nature Medicine,* 5**,** 340-343.

GURUPRASAD, K., REDDY, B. V. & PANDIT, M. W. (1990) Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting i*n vivo* stability of a protein from its primary sequence. *Protein Engineering,* 4**,** 155-61.

HAASE, S., CABRERA, A., LANGER, C., TREECK, M., STRUCK, N., HERRMANN, S., JANSEN, P. W., BRUCHHAUS, I., BACHMANN, A., DIAS, S., COWMAN, A. F., STUNNENBERG, H. G., SPIELMANN, T. & GILBERGER, T. W. (2008) Characterization of a conserved rhoptry-associated leucine zipper-like protein in the malaria parasite *Plasmodium falciparum*. *Infection and Immunity,* 76**,** 879-887.

HAEGGSTRÖM, M., KIRONDE, F., BERZINS, K., CHEN, Q., WAHLGREN, M. & FERNANDEZ, V. (2004) Common trafficking pathway for variant antigens destined

for the surface of the *Plasmodium falciparum*-infected erythrocyte. *Molecular and Biochemical Parasitology,* 133**,** 1-14.

HALL, N. & GARDNER, M. J. (2004) The Genome of *Plasmodium falciparum*. IN WATERS, A. P. & JANSE, C. J. (Eds.) *Malaria parasites, Genomes and Molecular Biology.* Caister Academic press, UK.

HALL, N., KARRAS, M., RAINE, J. D., CARLTON, J. M., KOOIJ, T. W. A., BERRIMAN, M., FLORENS, L., JANSSEN, C. S., PAIN, A., CHRISTOPHIDES, G. K., JAMES, K., RUTHERFORD, K., HARRIS, B., HARRIS, D., CHURCHER, C., QUAIL, M. A., ORMOND, D., DOGGETT, J., TRUEMAN, H. E., MENDOZA, J., BIDWELL, S. L., RAJANDREAM, M. A., CARUCCI, D. J., YATES III, J. R., KAFATOS, F. C., JANSE, C. J., BARRELL, B., TURNER, C. M. R., WATERS, A. P. & SINDEN, R. E. (2005) A comprehensive survey of the *Plasmodium* life cycle by genomic, transcriptomic, and proteomic analyses. *Science,* 307**,** 82-86.

HALL, T. A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series,* 41**,** 95-98.

HANSSEN, E., HAWTHORNE, P., DIXON, M. W. A., TRENHOLME, K. R., MCMILLAN, P. J., SPIELMANN, T., GARDINER, D. L. & TILLEY, L. (2008) Targeted mutagenesis of the ring-exported protein-1 of *Plasmodium falciparum* disrupts the architecture of Maurer's cleft organelles. *Molecular Microbiology,* 69**,** 938-953.

HAWKING, F., GAMMAGE, K. & WORMS, M. J. (1972) The asexual and sexual circadian rhythms of *Plasmodium vinckei chabaudi*, of *P. berghei* and of *P. gallinaceum Parasitology,* 65**,** 189-201.

HAWTHORNE, P. L., TRENHOLME, K. R., SKINNER-ADAMS, T. S., SPIELMANN, T., FISCHER, K., DIXON, M. W. A., ORTEGA, M. R., ANDERSON, K. L., KEMP, D. J. & GARDINER, D. L. (2004) A novel *Plasmodium falciparum* ring stage protein, REX, is located in Maurer's clefts. *Molecular and Biochemical Parasitology,* 136**,** 181-189.

HAYAKAWA, T., CULLETON, R., OTANI, H., HORII, T. & TANABE, K. (2008) Big bang in the evolution of extant malaria parasites. *Molecular Biology and Evolution,* 25**,** 2233-2239.

HEALER, J., MCGUINNESS, D., HOPCROFT, P., HALEY, S., CARTER, R. & RILEY, E. (1997) Complement-mediated lysis of *Plasmodium falciparum* gametes by malaria-immune human sera is associated with antibodies to the gamete surface antigen Pfs230. *Infection and Immunity,* 65**,** 3017-3023.

HEDDINI, A., PETTERSSON, F., KAI, O., SHAFI, J., OBIERO, J., CHEN, Q., BARRAGAN, A., WAHLGREN, M. & MARSH, K. (2001) Fresh isolates from children with severe *Plasmodium falciparum* malaria bind to multiple receptors. *Infection and Immunity,* 69**,** 5849-5856.

HELMBY, H., CAVELIER, L., PETTERSSON, U. & WAHLGREN, M. (1993) Rosetting *Plasmodium falciparum*-infected erythrocytes express unique strain-specific antigens on their surface. *Infection and Immunity,* 61**,** 284-288.

HENSMANN, M., LI, C., MOSS, C., LINDO, V., GREER, F., WATTS, C., OGUN, S. A., HOLDER, A. A. & LANGHORNE, J. (2004) Disulfide bonds in merozoite surface protein 1 of the malaria parasite impede efficient antigen processing and affect the *in vivo* antibody response. *European Journal of Immunology,* 34**,** 639-648.

HILLER, N. L., BHATTACHARJEE, S., VAN OOIJ, C., LIOLIOS, K., HARRISON, T., LOPEZ-ESTRAÑO, C. & HALDAR, K. (2004) A host-targeting signal in virulence proteins reveals a secretome in malarial infection. *Science,* 306**,** 1934-1937.

HIRUNPETCHARAT, C., TIAN, J. H., KASLOW, D. C., VAN ROOIJEN, N., KUMAR, S., BERZOFSKY, J. A., MILLER, L. H. & GOOD, M. F. (1997) Complete protective immunity induced in mice by immunization with the 19-kilodalton carboxyl-terminal fragment of the merozoite surface protein-1 (MSP1[19]) of *Plasmodium yoelii* expressed in *Saccharomyces cerevisiae*: correlation of protection with antigen-specific

antibody titer, but not with effector CD4+ T cells. *Journal of Immunology,* 159**,** 3400–3411.

HODDER, A. N., CREWTHER, P. E. & ANDERS, R. F. (2001) Specificity of the protective antibody response to apical membrane antigen 1. *Infection and Immunity,* 69**,** 3286–3294.

HOFFMAN, S. L., GOH, L. M., LUKE, T. C., SCHNEIDER, I., LE, T. P., DOOLAN, D. L., SACCI, J., DE LA VEGA, P., DOWLER, M., PAUL, C., GORDON, D. M., STOUTE, J. A., CHURCH, L. W., SEDEGAH, M., HEPPNER, D. G., BALLOU, W. R. & L., R. T. (2002) Protection of humans against malaria by immunization with radiation-attenuated *Plasmodium falciparum* sporozoites. *Journal of Infectious Diseases,* 185**,** 1155-1164.

HOFFMAN, S. L., OSTER, C. N. & PLOWE, C. V. (1987) Naturally acquired antibodies to sporozoites do not prevent malaria: Vaccine development implications. *Science,* 237**,** 639-642.

HOGSTRAND, K. & BOHME, J. (1999) Gene conversion can create new MHC alleles. *Immunological Reviews,* 167**,** 305-317.

HORATA, N., KALAMBAHETI, T., CRAIG, A. & KHUSMITH, S. (2009) Sequence variation of *Pf*EMP1-DBL alpha in association with rosette formation in *Plasmodium falciparum* isolates causing severe and uncomplicated malaria. *Malaria Journal,* 8.

HORROCKS, P., KYES, S., PINCHES, R., CHRISTODOULOU, Z. & NEWBOLD, C. (2004a) Transcription of a subtelomerically located *var* gene variant in *Plasmodium falciparum* appears to require the truncation of an adjacent *var* gene. *Molecular and Biochemical Parasitology,* 134**,** 193-199.

HORROCKS, P., PINCHES, R., CHRISTODOULOU, Z., KYES, S. A. & NEWBOLD, C. I. (2004b) Variable *var* transition rates underlie antigenic variation in malaria. *Proceedings of the National Academy of Sciences of the United States of America,* 101**,** 11129-34.

HOWARD, R. J., BARNWELL, J. W., ROCK, E. P., NEEQUAYE, J., OFORI-ADJEI, D., MALOY, W. L., LYON, J. A. & SAUL, A. (1988) Two approximately 300 kilodalton *Plasmodium falciparum* proteins at the surface membrane of infected erythrocytes. *Molecular and Biochemical Parasitology,* 27**,** 207-224.

HOWITT, C. A., WILLNSKI, D., LLINÁS, M., TEMPLETON, T. J., DZLKOWSKI, R. & DEITSCH, K. W. (2009) Clonally variant gene families in *Plasmodium falciparum* share a common activation factor. *Molecular Microbiology,* 73**,** 1171-1185.

HUEBERT, R. C., JAGAVELU, K., LIEBL, A. F., HUANG, B. Q., SPLINTER, P. L., LARUSSO, N. F., URRUTIA, R. A. & SHAH, V. H. (2010) Immortalized liver endothelial cells: a cell culture model for studies of motility and angiogenesis. *Laboratory Investigation,* 90**,** 1770-1781.

HUGHES, K. R., BIAGINI, G. A. & CRAIG, A. G. (2010) Continued cytoadherence of *Plasmodium falciparum* infected red blood cells after antimalarial treatment. *Molecular and Biochemical Parasitology,* 169**,** 71-78.

HUGUES, S., FETLER, L., BONIFAZ, L., HELFT, J., AMBLARD, F. & AMIGORENA, S. (2004) Distinct T cell dynamics in lymph nodes during the induction of tolerance and immunity. *Nature Immunology,* 5**,** 1235–1242.

HUSON, D. H. (1998) SplitsTree: A program for analyzing and visualizing evolutionary data. *Bioinformatics,* 14**,** 68-73.

HUSON, D. H. & BRYANT, D. (2006) Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution,* 23**,** 254-267.

HUSSAIN REED, Z., KIENY, M., P., ENGERS, H., FRIEDE, M., CHANG, S., LONGACRE, S., MALHOTRA, P., PAN, W. & LONG, C. (2009) Comparison of immunogenicity of five MSP1-based malaria vaccine candidate antigens in rabbits. *Vaccine,* 27**,** 1651-1660.

HVIID, L. (2010) The role of *Plasmodium falciparum* variant surface antigens in protective immunity and vaccine development. *Human Vaccines,* 6**,** 84-9.

JACKSON, A. P. (2007) Tandem gene arrays in *Trypanosoma brucei*: Comparative phylogenomic analysis of duplicate sequence variation. *BMC Evolutionary Biology* 7**,** 54.

JALAH, R., SARIN, R., SUD, N., ALAM, M. T., PARIKH, N., DAS, T. K. & SHARMA, Y. D. (2005) Identification, expression, localization and serological characterization of a tryptophan-rich antigen from the human malaria parasite *Plasmodium vivax*. *Molecular and Biochemical Parasitology,* 142**,** 158-169.

JANES, J. H., WANG, C. P., LEVIN-EDENS, E., VIGAN-WOMAS, I., GUILLOTTE, M., MELCHER, M., MERCEREAU-PUIJALON, O. & SMITH, J. D. (2011) Investigating the Host Binding Signature on the *Plasmodium falciparum Pf*EMP1 Protein Family. *PLoS Pathogens,* 7**,** e1002032.

JANEWAY, C. (2007) Immunobiology. IN MURPHY, K. M., TRAVERS, P. & WALPORT, M. (Eds.) Seventh ed.

JANSSEN, C. S., BARRETT, M. P., LAWSON, D., QUAIL, M. A., HARRIS, D., BOWMAN, S., PHILLIPS, R. S. & TURNER, C. M. R. (2001) Gene discovery in *Plasmodium chabaudi* by genome survey sequencing. *Molecular and Biochemical Parasitology,* 113**,** 251-260.

JANSSEN, C. S., BARRETT, M. P., TURNER, C. M. R. & PHILLIPS, S. R. (2002) A large gene family for putative variant antigens shared by human and rodent malaria parasites. *Proceedings of the Royal Society - Biological Sciences (Series B),* 269**,** 431-436.

JANSSEN, C. S., PHILLIPS, R. S., TURNER, M. R. & BARRET, M. P. (2004) *Plasmodium* interspersed repeats: The major multigene superfamily of malaria parasites. *Nucleic Acids Research,* 32**,** 5712-5720.

JARRA, W. & BROWN, K. N. (1989) Protective immunity to malaria: studies with cloned lines of rodent malaria in CBA/Ca mice. IV. The specificity of mechanisms resulting in crisis and resolution of the primary acute phase parasitaemia of *Plasmodium chabaudi chabaudi* and *P. yoelii yoelii*. . *Parasite Immunology,* 11**,** 1-13.

JARRA, W., HILLS, L. A., MARCH, J. C. & BROWN, K. N. (1986) Protective immunity to malaria. Studies with cloned lines of *Plasmodium chabaudi chabaudi* and *P. berghei* in CBA/Ca mice. II. The effectiveness and inter- or intra-species specificity of the passive transfer of immunity with serum. . *Parasite Immunology,* 8**,** 239-254.

JENKINS, M., K., KHORUTS, A., INGULLI, E., MUELLER, D. L., MCSORLEY, S. J., REINHARDT, R. L., ITANO, A. & PAPE, K. A. (2001) *In vivo* activation of antigen specific CD4 T Cells. *Annual Review of Immunology,* 19**,** 23-45.

JENNINGS, R. M., DE SOUZA, J. B., TODD, J. E., ARMSTRONG, M., FLANAGAN, K. L., RILEY, E. M. & DOHERTY, J. F. (2006) Imported *Plasmodium falciparum* malaria: Are patients originating from disease-endemic areas less likely to develop severe disease? A prospective, observational study. *American Journal of Tropical Medicine and Hygiene,* 75**,** 1195-1199.

JENSEN, A. T., MAGISTRADO, P., SHARP, S., JOERGENSEN, L., LAVSTSEN, T., CHIUCCHIUINI, A., SALANTI, A., VESTERGAARD, L. S., LUSINGU, J. P., HERMSEN, R., SAUERWEIN, R., CHRISTENSEN, J., NIELSEN, M. A., HVIID, L., SUTHERLAND, C., STAALSOE, T. & THEANDER, T. G. (2004) *Plasmodium falciparum* associated with severe childhood malaria preferentially expresses *Pf*EMP1 encoded by group A *var* genes. *Journal of Experimental Medicine,* 199**,** 1179-90.

JOANNIN, N. (2010) Antigenic variation in *Plasmodium falciparum*: Understanding the RIFIN protein family. *Karolinska Institutet.* Stockholm.

JOANNIN, N., ABHIMAN, S., SONNHAMMER, E. L. & WAHLGREN, M. (2008) Sub-grouping and sub-functionalization of the RIFIN multi-copy protein family. *BMC Genomics,* 9.

JOANNIN, N., KALLBERG, Y., WAHLGREN, M. & PERSSON, B. (2010) RSpred, a set of Hidden Markov Models to detect and classify the RIFIN and STEVOR proteins of *Plasmodium falciparum*. *BMC Genomics,* 12**,** 119.

JOERGENSEN, L., BENGTSSON, D. C., BENGTSSON, A., RONANDER, E., BERGER, S. S., TURNER, L., DALGAARD, M. B., CHAM, G. K. K., VICTOR, M. E., LAVSTSEN, T., THEANDER, T. G., ARNOT, D. E. & JENSEN, A. T. R. (2010) Surface co-expression of two different *Pf*EMP1 antigens on single *Plasmodium falciparum* -infected erythrocytes facilitates binding to ICAM1 and PECAM1. *PLoS Pathogens,* 6.

JOERGENSEN, L., TURNER, L., MAGISTRADO, P., DAHLBÄCK, M. A., VESTERGAARD, L. S., LUSINGU, J. P., LEMNGE, M., SALANTI, A., THEANDER, T. G. & JENSEN, A. T. R. (2006) Limited cross-reactivity among domains of the *Plasmodium falciparum* clone 3D7 erythrocyte membrane protein 1 family. *Infection and Immunity,* 74**,** 6778-6784.

JOUVIN-MARCHE, E., GONCALVES MORGADO, M., LEGUERN, C., VOEGTLE, D., BONHOMME, F. & CAZENAVE, P. A. (1989) The mouse Igh-1a and Igh-1b H chain constant regions are derived from two distinct isotypic genes. *Immunogenetics,* 29**,** 92-97.

KAESTLI, M., COCKBURN, I. A., CORTÉS, A., BAEA, K., ROWE, J. A. & BECK, H. P. (2006) Virulence of malaria is associated with differential expression of *Plasmodium falciparum var* gene subgroups in a case-control study. *Journal of Infectious Diseases,* 193**,** 1567-1574.

KAVIRATNE, M., KHAN, S. M., JARRA, W. & PREISER, P. R. (2002) Small variant STEVOR antigen is uniquely located within Maurer's clefts in *Plasmodium falciparum*-infected red blood cells. *Eukaryotic Cell,* 1**,** 926-35.

KAWAMOTO, Y., KOJIMA, K., HITSUMOTO, Y., OKADA, H., HOLERS, V. M. & MIYAMA, A. (1997) The serum resistance of malaria-infected erythrocytes. *Immunology,* 91**,** 7-12.

KEEN, J., SERGHIDES, L., AYI, K., PATEL, S. N., AYISI, J., VAN EIJK, A., STEKETEE, R., UDHAYAKUMAR, V. & KAIN, K. C. (2007) HIV impairs opsonic phagocytic clearance of pregnancy-associated malaria parasites. *PLoS Medicine,* 4**,** e181.

KHATTAB, A., BONOW, I., SCHREIBER, N., PETTER, M., SCHMETZ, C. & KLINKERT, M. Q. (2008) *Plasmodium falciparum* variant STEVOR antigens are expressed in merozoites and possibly associated with erythrocyte invasion. *Malaria Journal,* 7**,** 137.

KHATTAB, A. & MERI, S. (2011) Exposure of the *Plasmodium falciparum* clonally variant STEVOR proteins on the merozoite surface. *Malaria Journal,* 10.

KINA, T., IKUTA, K., TAKAYAMA, E., WADA, K., MAJUMDAR, A. S., WEISSMAN, I. L. & KATSURA, Y. (2000) The monoclonal antibody TER-119 recognizes a molecule associated with glycophorin A and specifically marks the late stages of murine erythroid lineage. *British Journal of Haematology,* 109**,** 280-287.

KINYANJUI, S., M., MWANGI, T., BULL, P., C. & AL., E. (2004a) Protection against clinical malaria by heterologous immunoglobulin G antibodies against malaria-infected erythrocyte variant surface antigens requires interaction with asymptomatic infections. *Journal of Infectious Diseases,* 190**,** 1527-1533.

KINYANJUI, S. M., CONWAY, D. J., LANAR, D. E. & MARSH, K. (2007) IgG antibody responses to *Plasmodium falciparum* merozoite antigens in Kenyan children have a short half-life. *Malaria Journal,* 6**,** 82-90.

KINYANJUI, S. M., HOWARD, T., WILLIAMS, T. N., BULL, P. C., NEWBOLD, C. I. & MARSH, K. (2004b) The use of cryopreserved mature trophozoites in assessing antibody recognition of variant surface antigens of *Plasmodium falciparum*-infected erythrocytes. *Journal of Immunological Methods,* 288**,** 9-18.

KITUA, A. Y., URASSA, H., WECHSLER, M., SMITH, T., VOUNATSOU, P., WEISS, N. A., ALONSO, P. L. & TANNER, M. (1999) Antibodies against *Plasmodium falciparum* vaccine candidates in infants in an area of intense and perennial transmission: relationships with clinical malaria and with entomological inoculation rates. *Parasite immunology,* 21**,** 307-317.

KNUDSEN, B. & MIYAMOTO, M. M. (2001) A likelihood ratio test for evolutionary rate shifts and functional divergence among proteins. *Proceedings of the National Academy of Sciences of the United States of America,* 98**,** 14512-14517.

KOCKEN, C. H. M., DUBBELD, M. A., VAN DER WEL, A., PRONK, J. T., WATERS, A. P., LANGERMANS, J. A. M. & THOMAS, A. W. (1999) High-level expression of *Plasmodium vivax* apical membrane antigen 1 (AMA-1) in *Pichia pastoris*: Strong immunogenicity in *Macaca mulatta* immunized with *P. vivax* AMA-1 and adjuvant SBAS2. *Infection and Immunity,* 67**,** 43-49.

KOCKEN, C. H. M., WITHERS-MARTINEZ, C., DUBBELD, M. A., VAN DER WEL, A., HACKETT, F., BLACKMAN, M. J. & THOMAS, A. W. (2002) High-level expression of the malaria blood-stage vaccine candidate *Plasmodium falciparum* apical membrane antigen 1 and induction of antibodies that inhibit erythrocyte invasion. *Infection and Immunity,* 70**,** 4471-4476.

KOUTZ, P. J., DAVIS, G. R., STILLMAN, C., BARRINGER, K., CREGG, J. M. & THILL, G. (1989) Structural Comparison of the *Pichia pastoris* Alcohol Oxidase Genes. *Yeast,* 5**,** 167-177.

KRAEMER, S. M., KYES, S. A., AGGARWAL, G., SPRINGER, A. L., NELSON, S. O., CHRISTODOULOU, Z., SMITH, L. M., WANG, W., LEVIN, E., NEWBOLD, C. I., MYLER, P. J. & SMITH, J. D. (2007) Patterns of gene recombination shape *var* gene repertoires in *Plasmodium falciparum*: Comparisons of geographically diverse isolates. *BMC Genomics,* 8.

KRAEMER, S. M. & SMITH, J. D. (2003) Evidence for the importance of genetic structuring to the structural and functional specialization of the *Plasmodium falciparum var* gene family. *Molecular Microbiology,* 50**,** 1527-38.

KREBS, H. & EGGLESTON, L. V. (1940) The oxidation of pyruvate in pigeon breast muscle. *Biochemical Journal,* 34**,** 442-459.

KROTOSKI, W. A., COLLINS, W. E. & BRAY, R. S. (1982) Demonstration of hypnozoites in sporozoite-transmitted *Plasmodium vivax* infection. *American Journal of Tropical Medicine and Hygiene,* 31**,** 1291-1293.

KUDLA, G., HELWAK, A. & LIPINSKI, L. (2004) Gene conversion and GC-content evolution in mammalian Hsp70. *Molecular Biology and Evolution,* 21**,** 1438-1444.

KUHN, Y., SANCHEZ, C. P., AYOUB, D., SARIDAKI, T., VAN DORSSELAER, A. & LANZER, M. (2010) Trafficking of the Phosphoprotein *Pf*CRT to the Digestive Vacuolar Membrane in *Plasmodium falciparum*. *Traffic,* 11**,** 236-249.

KUMAR, S., TAMURA, K. & NEI, M. (2004) MEGA3: Integrated Software for Molecular Evolutionary Analysis and Sequence Alignment. *Briefings in Bioinformatics,* 5**,** 150-163.

KURTIS, J. D., MTALIB, R., ONYANGO, F. K. & DUFFY, P. E. (2001) Human resistance to *Plasmodium falciparum* increases during puberty and is predicted by dehydroepiandrosterone sulfate levels. *Infection and Immunity,* 69**,** 123-128.

KURTZHALS, J. A. L., ADABAYERI, V., GOKA, B. Q., AKANMORI, B. D., OLIVER-COMMEY, J. O., NKRUMAH, F. K., BEHR, C. & HVIID, L. (1998) Low plasma concentrations of interleukin-10 in severe malarial anaemia compared with cerebral and uncomplicated malaria. *Lancet,* 351**,** 1768-1772.

KURZ, M., COWIESON, N. P., ROBIN, G., HUME, D. A., MARTIN, J. L., KOBE, B. & LISTWAN, P. (2006) Incorporating a TEV cleavage site reduces the solubility of nine recombinant mouse proteins. *Protein Expression and Purification,* 50**,** 68-73.

KWIATKOWSKI, D. (1990) Tumour necrosis factor, fever and fatality in *P. falciparum* malaria. *Immunology Letters,* 25**,** 213-216.

KYES, S., HORROCKS, P. & NEWBOLD, C. (2001) Antigenic variation at the infected red cell surface in malaria. *Annual Review of Microbiology,* 55**,** 673-707.

KYES, S. A., KRAEMER, S. M. & SMITH, J. D. (2007) Antigenic variation in *Plasmodium falciparum*: Gene organization and regulation of the *var* multigene family. *Eukaryotic Cell,* 6**,** 1511-1520.

KYES, S. A., ROWE, J. A., KRIEK, N. & NEWBOLD, C. I. (1999) Rifins: a second family of clonally variant proteins expressed on the surface of red cells infected with *Plasmodium falciparum*. *Proceedings of the National Academy of Sciences of the United States of America,* 96**,** 9333-8.

KYRIACOU, H. M., STONE, G. N., CHALLIS, R. J., RAZA, A., LYKE, K. E., THERA, M. A., KONÉ, A. K., DOUMBO, O. K., PLOWE, C. V. & ROWE, J. A. (2006) Differential *var* gene transcription in *Plasmodium falciparum* isolates from patients with cerebral malaria compared to hyperparasitaemia. *Molecular and Biochemical Parasitology,* 150**,** 211-218.

KYTE, J. & DOOLITTLE, R. F. (1982) A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology,* 157**,** 105-132.

LAMB, T. J., BROWN, D. E., POTOCNIK, A. J. & LANGHORNE, J. (2006) Insights into the immunopathogenesis of malaria using mouse models. *Expert Reviews in Molecular Medicine,* 8**,** 1-22.

LANDAU, I. (1965) Description of *Plasmodium chabaudi* n. sp. parasite of African rodents. *Comptes Rendus Hebdomadaires des Seances de l Academie des Sciences. D,* 260**,** 3758-61.

LANDAU, I. & BOULARD, Y. (1978) Life cycles and morphology. *Rodent Malaria***,** 53-84.

LANGHORNE, J. (1989) The role of CD4+ T-cells in the immune response to *Plasmodium chabaudi*. *Parasitology Today,* 5**,** 362-364.

LANGHORNE, J., ALBANO, F. R., HENSMANN, M., SANNI, L., CADMAN, E., VOISINE, C. & SPONAAS, A. M. (2004) Dendritic cells, pro-inflammatory responses, and antigen presentation in a rodent malaria infection. *Immunological Reviews,* 201**,** 35-47.

LANGHORNE, J., GILLARD, S., SIMON, B., SLADE, S. & EICHMANN, K. (1989) Frequencies of CD4+ T cells reactive with *Plasmodium chabaudi chabaudi*: distinct response kinetics for cells with T(h1) and T(h2) characteristics during infection. *International Immunology,* 1**,** 416-424.

LANGHORNE, J., NDUNGU, F. M., SPONAAS, A. M. & MARSH, K. (2008) Immunity to malaria: More questions than answers. *Nature Immunology,* 9**,** 725-732.

LAU, O. S., NG, D. W. K., CHAN, W. W., CHANG, S. P. & SUN, S. S. (2010) Production of the 42-kDa fragment of *Plasmodium falciparum* merozoite surface protein 1, a leading malaria vaccine antigen, in *Arabidopsis thaliana* seeds. *Plant Biotechnology Journal,* 8**,** 994-1004.

LAVAZEC, C., SANYAL, S. & TEMPLETON, T. J. (2006) Hypervariability within the Rifin, Stevor and *Pf*mc-2TM superfamilies in *Plasmodium falciparum*. *Nucleic Acids Research,* 34**,** 6696-6707.

LAVAZEC, C., SANYAL, S. & TEMPLETON, T. J. (2007) Expression switching in the stevor and *Pf*mc-2TM superfamilies in *Plasmodium falciparum*. *Molecular Microbiology,* 64**,** 1621-34.

LAVSTSEN, T., MAGISTRADO, P., HERMSEN, C. C., SALANTI, A., JENSEN, A. T. R., SAUERWEIN, R., HVIID, L., THEANDER, T. G. & STAALSOE, T. (2005) Expression of *Plasmodium falciparum* erythrocyte membrane protein 1 in experimentally infected humans. *Malaria Journal,* 4.

LE ROCH, K. G., JOHNSON, J. R., FLORENS, L., ZHOU, Y., SANTROSYAN, A., GRAINGER, M., YAN, S. F., WILLIAMSON, K. C., HOLDER, A. A., CARUCCI, D. J., YATES III, J. R. & WINZELER, E. A. (2004) Global analysis of transcript and protein levels across the *Plasmodium falciparum* life cycle. *Genome Research,* 14**,** 2308-2318.

LE ROCH, K. G., ZHOU, Y., BLAIR, P. L., GRAINGER, M., MOCH, J. K., HAYNES, J. D., DE LA VEGA, P., HOLDER, A. A., BATALOV, S., CARUCCI, D. J. & WINZELER, E. A. (2003) Discovery of gene function by expression profiling of the malaria parasite life cycle. *Science,* 301**,** 1503-1508.

LEECH, J. H., BARNWELL, J. W., MILLER, L. H. & HOWARD, R. J. (1984) Identification of a strain-specific malarial antigen exposed on the surface of *Plasmodium falciparum*-infected erythrocytes. *Journal of Experimental Medicine,* 159**,** 1567-75.

LI, C., CORRALIZA, I. & LANGHORNE, J. (1999) A defect in interleukin-10 leads to enhanced malarial disease in *Plasmodium chabaudi chabaudi* infection in mice. *Infection and Immunity,* 67**,** 4435-4442.

LIEW, K. J. L., HU, G., BOZDECH, Z. & PETER, P. R. (2010) Defining species specific genome differences in malaria parasites. *BMC Genomics,* 11**,** 128.

LIN, K. W. & YAN, J. (2008) Endings in the middle: Current knowledge of interstitial telomeric sequences. *Mutation Research - Reviews in Mutation Research,* 658**,** 95-110.

LING, I., OGUN, S. & HOLDER, A. A. (1994) Immunisation against malaria with a recombinant protein. *Parasite Immunology,* 16**,** 63-67.

LING, I. T., OGUN, S. A., MOMIN, P., RICHARDS, R. L., GARCON, N., COHEN, J., BALLOU, W. R. & HOLDER, A. A. (1997) Immunization against the murine malaria parasite *Plasmodium yoelii* using a recombinant protein with adjuvants developed for clinical use. *Vaccine,* 15.

LLINÁS, M., BOZDECH, Z., WONG, E. D., ADAI, A. T. & DERISI, J. L. (2006) Comparative whole genome transcriptome analysis of three *Plasmodium falciparum* strains. *Nucleic Acids Research,* 34**,** 1166-1173.

LOPEZ-RUBIO, J. J., GONTIJO, A. M., NUNES, M. C., ISSAR, N., HERNANDEZ RIVAS, R. & SCHERF, A. (2007) 5' flanking region of *var* genes nucleate histone modification patterns linked to phenotypic inheritance of virulence traits in malaria parasites. *Molecular Microbiology,* 66**,** 1296-1305.

LORD, R., JONES, G. L., SPENCER, L. & SAUL, A. (1993) Mice immunized with a synthetic peptide construct corresponding to an epitope present on a *Plasmodium falciparum* antigen are protected against *Plasmodium chabaudi* challenge. *Parasite Immunology,* 15**,** 613-618.

LOVEGROVE, F. E., PEÑA-CASTILLO, L., MOHAMMAD, N., LILES, W. C., HUGHES, T. R. & KAIN, K. C. (2006) Simultaneous host and parasite expression profiling identifies tissue-specific transcriptional programs associated with susceptibility or resistance to experimental cerebral malaria. *BMC Genomics,* 7.

LOZANO, J. M., LESMES, L. P., GALLEGO, G. M. & PATARROYO, M. E. (2011) Protection against malaria is conferred by passive transferring rabbit F(ab)(2)' antibody fragments, induced by *Plasmodium falciparum* MSP-1 site-directed designed pseudopeptide-BSA conjugates assessed in a rodent model. *Molecular Immunology,* 48**,** 657-69.

LUTY, A. J., LELL, B., SCHMIDT-OTT, R., LEHMAN, L. G., LUCKNER, D., GREVE, B., MATOUSEK, P., HERBICH, K., SCHMID, D., MIGOT-NABIAS, F., DELORON, P., NUSSENZWEIG, R. S. & KREMSNER, P. G. (1999) Interferon-gamma responses are associated with resistance to reinfection with *Plasmodium falciparum* in young African children. *Journal of Infectious Diseases,* 179**,** 980-988.

MACKINNON, M. J., LI, J., MOK, S., KORTOK, M. M., MARSH, K., PREISER, P. R. & BOZDECH, Z. (2009) Comparative transcriptional and genomic analysis of *Plasmodium falciparum* field isolates. *PLoS Pathogens,* 5.

MACKINNON, M. J., MWANGI, T. W., SNOW, R. W., MARSH, K. & WILLIAMS, T. N. (2005) Heritability of malaria in Africa. *PLoS Medicine,* e340.

MACKINNON, M. J., WALKER, P. R. & ROWE, J. A. (2002) *Plasmodium chabaudi*: Rosetting in a rodent malaria model *Experimental Parasitology,* 101**,** 121-128.

MACKINTOSH, C. L., BEESON, J. G. & MARSH, K. (2004) Clinical features and pathogenesis of severe malaria. *Trends in Parasitology,* 20**,** 597-603.

MACKINTOSH, C. L., CHRISTODOULOU, Z., MWANGI, T. W., KORTOK, M., PINCHES, R., WILLIAMS, T. N., MARSH, K. & NEWBOLD, C. I. (2008) Acquisition of naturally occurring antibody responses to recombinant protein domains of *Plasmodium falciparum* erythrocyte membrane protein 1. *Malaria Journal,* 7**,** 155.

MACPHERSON, G. G., WARRELL, M. J. & WHITE, N. J. (1985) Human cerebral malaria. A quantitative ultrastructural analysis of parasitized erythrocyte sequestration. *American Journal of Pathology,* 119**,** 385-401.

MAGISTRADO, P. A., LUSINGU, J., VESTERGAARD, L. S., LEMNGE, M., LAVSTSEN, T., TURNER, L., HVIID, L., JENSEN, A. T. & THEANDER, T. G. (2007) Immunoglobulin G antibody reactivity to a group A *Plasmodium falciparum* erythrocyte membrane protein 1 and protection from *P. falciparum* malaria. *Infection and Immunity,* 75**,** 2415-20.

MAITLAND, K., LEVIN, M., ENGLISH, M., MITHWANI, S., PESHU, N., MARSH, K. & NEWTON, C. R. J. C. (2003) Severe *P. falciparum* malaria in Kenyan children: Evidence for hypovolaemia. *QJM - Monthly Journal of the Association of Physicians,* 96**,** 427-434.

MAIZELS, N. (2005) Immunoglobulin gene diversification. *Annual Review of Genetics,* 39**,** 23-46.

MARSH, K., FORSTER, D., WARUIRU, C., MWANGI, I., WINSTANLEY, M., MARSH, V., NEWTON, C., WINSTANLEY, P., WARN, P., PESHU, N., PASVOL, G. & SNOW, R. (1995) Indicators of life-threatening malaria in African children. *New England Journal of Medicine,* 332**,** 1399-1404.

MARSH, K. & HOWARD, R. J. (1986) Antigens induced on erythrocytes by *P. falciparum*: Expression of diverse and conserved determinants. *Science,* 231**,** 150-153.

MARSH, K. & KINYANJUI, S. (2006) Immune effector mechanisms in malaria. *Parasite Immunology,* 28**,** 51-60.

MARSH, K., OTOO, L., HAYES, R. J., CARSON, D. C. & GREENWOOD, B. M. (1989) Antibodies to blood stage antigens of *Plasmodium falciparum* in rural Gambians and their relation to protection against infection. *Transactions of the Royal Society of Tropical Medicine and Hygiene,* 83**,** 293-303.

MARSH, K. & SNOW, R. W. (1997) Host-parasite interaction and morbidity in malaria endemic areas. *Philosophical Transactions of the Royal Society B: Biological Sciences,* 352**,** 1385-1394.

MARTI, M., GOOD, R. T., RUG, M., KNUEPFER, E. & COWMAN, A. F. (2004) Targeting malaria virulence and remodeling proteins to the host erythrocyte. *Science,* 306**,** 1930-1933.

MARTIN, D. & RYBICKI, E. (2000) RDP: detection of recombination amongst aligned sequences. *Bioinformatics,* 16**,** 562-563.

MARTIN-JAULAR, L., FERRER, M., CALVO, M., ROSANAS-URGELL, A., KALKO, S., GRAEWE, S., SORIA, G., CORTADELLAS, N., ORDI, J., PLANAS, A., BURNS, J., HEUSSLER, V. & DEL PORTILLO, H. A. (2011) Strain-specific spleen remodelling in *Plasmodium yoelii* infections in Balb/c mice facilitates adherence and spleen macrophage-clearance escape. *Cellular Microbiology,* 13**,** 109-122.

MARTINSOHN, J. T., SOUSA, A. B., GUETHLEIN, L. A. & HOWARD, J. C. (1999) The gene conversion hypothesis of MHC evolution: a review. *Immunogenetics,* 50**,** 168-200.

MATSUMOTO, Y., AIKAWA, M. & BARNWELL, J. W. (1988) Immunoelectron microscopic localization of *vivax* malaria antigens to the clefts and caveola-vesicle complexes of infected erythrocytes. *American Journal of Tropical Medicine and Hygiene,* 39**,** 317-322.

MAUDE, R. J., BEARE, N. A. V., SAYEED, A. A., CHANG, C. C., CHARUNWATTHANA, P., FAIZ, M. A., HOSSAIN, A., YUNUS, E. B., HOQUE, M. G., HASAN, M. U., WHITE, N. J., DAY, N. P. J. & DONDORP, A. M. (2009a) The spectrum of retinopathy in adults with *Plasmodium falciparum* malaria. *Transactions of the Royal Society of Tropical Medicine and Hygiene,* 103**,** 665-671.

MAUDE, R. J., DONDORP, A. M., SAYEED, A. A., DAY, N. P. J., WHITE, N. J. & BEARE, N. A. V. (2009b) The eye in cerebral malaria: what can it teach us? *Transactions of the Royal Society of Tropical Medicine and Hygiene,* 103**,** 661-664.

MCCORMICK, C. J., CRAIG, A., ROBERTS, D., NEWBOLD, C. I. & BERENDT, A. R. (1997) Intercellular adhesion molecule-1 and CD36 synergize to mediate adherence of *Plasmodium falciparum*-infected erythrocytes to cultured human microvascular endothelial cells. *Journal of Clinical Investigation,* 100**,** 2521-2529.

MCGARRY, M. P., PROTHEROE, C. A. & LEE, J. J. (2010) *Mouse Hematology: A Laboratory Manual*, Cold Spring Harbor Laboratory Press.

MCGREGOR, I. A. & CARRINGTON, S. (1963) Treatment of East African *P. falciparum* malaria with West African gamma-globulin. *Transactions of the Royal Society of Tropical Medicine and Hygiene,* 57**,** 170-175.

MCINERNEY, J. O., LITTLEWOOD, D. T. & CREEVEY, C. J. (2003) Detecting adaptive molecular evolution: additional tools for the parasitologist. *Advances in Parasitology,* 54**,** 359-79.

MCKEAN, P. G., O'DEA, K. P. & BROWN, K. N. (1993) Nucleotide sequence analysis and epitope mapping of the merozoite surface protein 1 from *Plasmodium chabaudi chabaudi* AS. *Molecular and Biochemical Parasitology,* 62**,** 199-209.

MCKENZIE, F. E., SMITH, D. L., O'MEARA, W. P. & RILEY, E. M. (2008) Chapter 1 Strain Theory of Malaria. The First 50 Years. *Advances in Parasitology,* 66**,** 1-46.

MCLEAN, S. A., PEARSON, C. D. & PHILLIPS, R. S. (1982) *Plasmodium chabaudi*: Relationship between the occurrence of recrudescent parasitaemias in mice and the effective levels of acquired immunity. *Experimental Parasitology,* 54**,** 213-221.

MCLEAN, S. A., PEARSON, C. D. & PHILLIPS, R. S. (1986) Antigenic variation in *Plasmodium chabaudi*: Analysis of parent and variant populations by cloning. *Parasite Immunology,* 8**,** 415-424.

MCROBERT, L., PREISER, P., SHARP, S., JARRA, W., KAVIRATNE, M., TAYLOR, M. C., RENIA, L. & SUTHERLAND, C. J. (2004) Distinct trafficking and localization of STEVOR proteins in three stages of the *Plasmodium falciparum* life cycle. *Infection and Immunity,* 72**,** 6597-6602.

MEDANA, I. M. & TURNER, G. D. H. (2006) Human cerebral malaria and the blood-brain barrier. *International Journal for Parasitology,* 36**,** 555-568.

MEDING, S., J. & LANGHORNE, J. (1991) CD4+ T cells and B cells are necessary for the transfer of protective immunity to *Plasmodium chabaudi chabaudi*. *European Journal of Immunology,* 21**,** 1433-1438.

MEHLIN, C., BONI, E., BUCKNER, F. S., ENGEL, L., FEIST, T., GELB, M. H., HAJI, L., KIM, D., LIU, C., MUELLER, N., MYLER, P. J., REDDY, J. T., SAMPSON, J. N., SUBRAMANIAN, E., VAN VOORHIS, W. C., WORTHEY, E., ZUCKER, F. & HOL, W. G. J. (2006) Heterologous expression of proteins from *Plasmodium falciparum*: Results from 1000 genes. *Molecular and Biochemical Parasitology,* 148**,** 144-160.

MENENDEZ, C., FLEMING, A. F. & ALONSO, P. L. (2000) Malaria-related anaemia. *Parasitology Today,* 16**,** 469-476.

MERINO, E. F., FERNANDEZ-BECERRA, C., DURHAM, A. M., FERREIRA, J. E., TUMILASCI, V. F., D'ARC-NEVES, J., DA SILVA-NUNES, M., FERREIRA, M. U., WICKRAMARACHCHI, T., UDAGAMA-RANDENIYA, P., HANDUNNETTI, S. M. & DEL PORTILLO, H. A. (2006) Multi-character population study of the *vir* subtelomeric multigene superfamily of *Plasmodium vivax*, a major human malaria parasite. *Molecular and Biochemical Parasitology,* 149**,** 10-16.

METZGER, W. G., OKENU, D. M., CAVANAGH, D. R., ROBINSON, J. V., BOJANG, K. A., WEISS, H. A., MCBRIDE, J. S., GREENWOOD, B. M. & CONWAY, D. J. (2003) Serum IgG3 to the *Plasmodium falciparum* merozoite surface protein 2 is strongly associated with a reduced prospective risk of malaria. *Parasite Immunology,* 25**,** 307-312.

MICHELIN, A., BITTAME, A., BORDAT, Y., TRAVIER, L., MERCIER, C., DUBREMETZ, J. F. & LEBRUN, M. (2009) GRA12, a *Toxoplasma* dense granule protein associated with the intravacuolar membranous nanotubular network. *International Journal for Parasitology,* 39**,** 299-306.

MIGOT, F., CHOUGNET, C., RAHARIMALALA, L., ASTAGNEAU, P., LEPERS, J. P. & DELORON, P. (1993) Human immune responses to the *Plasmodium falciparum* ring-infected erythrocyte surface antigen (Pf155/RE SA) after a decrease in malaria transmission in Madagascar. *American Journal of Tropical Medicine and Hygiene,* 48**,** 432-439.

MILLER, L., H. , GOOD, M. F. & MILON, G. (1994) Malaria pathogenesis. *Science,* 264**,** 1878-1883.

MILLER, L. H. (1999) Evolution of the human genome under selective pressure from malaria: Applications for control. *Parasitologia,* 41**,** 77-82.

MILLER, L. H., AIKAWA, M. & DVORAK, J. A. (1975) Malaria (*Plasmodium knowlesi*) merozoites: Immunity and the surface coat. . *Journal of Immunology,* 114**,** 1237-1242.

MILLER, L. H., MASON, S. J., CLYDE, D. F. & MCGINNISS, M. H. (1976) The resistance factor to *Plasmodium vivax* in blacks. The Duffy-blood-group genotype, FyFy. *New England Jounal of Medicine,* 295**,** 302-304.

MILNE, I., LINDNER, D., BAYER, M., HUSMEIER, D., MCGUIRE, G., MARSHALL, D., F. & WRIGHT, F. (2009) TOPALi v2: a rich graphical interface for evolutionary analyses of multiple alignments on HPC clusters and multi-core desktops. *Bioinformatics,* 25**,** 126-127.

MILNE, I., WRIGHT, F., ROWE, G., MARSHALL, D. F., HUSMEIER, D. & MCGUIRE, G. (2004) TOPALi: Software for automatic identification of recombinant sequences within DNA multiple alignments. *Bioinformatics,* 20**,** 1806-7.

MIU, J., HUNT, N. H. & BALL, H. J. (2008) Predominance of interferon-related responses in the brain during murine malaria, as identified by microarray analysis. *Infection and Immunity,* 76**,** 1812-1824.

MO, M., HOOI, C. L., KOTAKA, M., NIANG, M., GAO, X., IYER, J. K., LESCAR, J. & PREISER, P. (2008) The C-terminal segment of the cysteine-rich interdomain of *Plasmodium falciparum* erythrocyte membrane protein 1 determines CD36 binding and elicits antibodies that inhibit adhesion of parasite-infected erythrocytes. *Infection and Immunity,* 76**,** 1837-1847.

MONTGOMERY, J., MPHANDE, F. A., BERRIMAN, M., PAIN, A., ROGERSON, S. J., TAYLOR, T. E., MOLYNEUX, M. E. & CRAIG, A. (2007) Differential *var* gene expression in the organs of patients dying of *falciparum* malaria. *Molecular Microbiology,* 65**,** 959-67.

MOONEN, B., COHEN, J. M., SNOW, R. W., SLUTSKER, L., DRAKELEY, C., SMITH, D. L., ABEYASINGHE, R. R., RODRIGUEZ, M. H., MAHARAJ, R., TANNER, M. & TARGETT, G. (2010) Operational strategies to achieve and maintain malaria elimination. *Lancet,* 376**,** 1592-1603.

MORGADO, M. G., CAM, P., GRIS-LIEBE, C., CAZENAVA, P. A. & JOUVIN-MARCHE, E. (1989) Further evidence that BALB/c and C57BL/6 γ2a genes originate from two distinct isotypes. *EMBO Journal,* 8**,** 3245-3251.

MORRISON, D. A. (2005) Networks in phylogenetic analysis: new tools for population biology. *International Journal for Parasitology,* 35**,** 567–582.

MOTA, M. M., BROWN, K. N., DO ROSÁRIO, V. E., HOLDER, A. A. & JARRA, W. (2001) Antibody recognition of rodent malaria parasite antigens exposed at the infected erythrocyte surface: Specificity of immunity generated in hyperimmune mice. *Infection and Immunity,* 69**,** 2535-2541.

MOTA, M. M., BROWN, K. N., HOLDER, A. A. & JARRA, W. (1998) Acute *Plasmodium chabaudi chabaudi* malaria infection induces antibodies which bind to the surfaces of parasitized erythrocytes and promote their phagocytosis by macrophages *in vitro*. *Infection and Immunity,* 66**,** 4080-4086.

MOTA, M. M., JARRA, W., HIRST, E., PATNAIK, P. K. & HOLDER, A. A. (2000) *Plasmodium chabaudi*-infected erythrocytes adhere to CD36 and bind to microvascular endothelial cells in an organ-specific way. *Infection and Immunity,* 68**,** 4135-4144.

MOXON, C. A., GRAU, G. E. & CRAIG, A. G. (2010) Malaria: Modification of the red blood cell and consequences in the human host. *British Journal of Haematology,* 154**,** 670-679.

MU, J., AWADALLA, P., DUAN, J., MCGEE, K. M., JOY, D. A., MCVEAN, G. A. T. & SU, X. Z. (2005) Recombination hotspots and population structure in *Plasmodium falciparum*. *PLoS Biology,* 3**,** e335.

MU, J., AWADALLA, P., DUAN, J., MCGEE, K. M., KEEBLER, J., SEYDEL, K., MCVEAN, G. A. T. & SU, X. Z. (2007) Genome-wide variation and identification of vaccine targets in the *Plasmodium falciparum* genome. *Nature Genetics,* 39**,** 126-130.

MUELLER, A. K., CAMARGO, N., KAISER, K., ANDORFER, C., FREVERT, U., MATUSCHEWSKI, K. & KAPPE, S. H. I. (2005) *Plasmodium* liver stage developmental arrest by depletion of a protein at the parasite–host interface. *Proceedings of the National Academy of Sciences of the United States of America,* 102**,** 3022–3027.

NARUM, D. L., OGUN, S. A., THOMAS, A. W. & HOLDER, A. A. (2000) Immunization with parasite-derived apical membrane antigen 1 or passive immunization with a specific monoclonal antibody protects BALB/c mice against lethal *Plasmodium yoelii yoelii* YM blood-stage infection. *Infection and Immunity,* 68**,** 2899-2906.

NARUM, D. L. & THOMAS, A. W. (1994) Differential localization of full-length and processed forms of PF83/AMA-1 an apical membrane antigen of *Plasmodium falciparum* merozoites. *Molecular and Biochemical Parasitology,* 67**,** 59-68.

NDUATI, E. W. & NG, D. H. L. (2010) Distinct kinetics of memory B-cell and plasma-cell responses in peripheral blood following a blood-stage *Plasmodium chabaudi* infection in mice. *PLoS One*.

NDUNGU, F., M. , CADMAN, E., T., COULCHER, J., NDUATI, E., COUPER, E., MACDONALD, D. W., NG, D. & LANGHORNE, J. (2009) Functional Memory B Cells and Long-Lived Plasma Cells Are Generated after a Single *Plasmodium chabaudi* Infection in Mice. . *PLoS Pathogens,* 5.

NDUNGU, F. M., SANNI, L., URBAN, B., STEPHENS, R., NEWBOLD, C. I., MARSH, K. & LANGHORNE, J. (2006) CD4 T cells from malaria-nonexposed individuals respond to the CD36-binding domain of *Plasmodium falciparum* erythrocyte membrane protein-1 via an MHC class II-TCR-independent pathway. *Journal of Immunology,* 176**,** 5504-5512.

NEBIE, I., DIARRA, A., OUEDRAOGO, A., SOULAMA, I., BOUGOUMA, E., TIONO, A., KONATE, A., CHILENGI, R., THEISEN, M., DODOO, D., REMARQUE, E., BOSOMPRAH, S., MILLIGAN, P. & SIRIMA, S. (2008) Humoral responses to *Plasmodium falciparum* blood-stage antigens and association with incidence of clinical malaria in children living in an area of seasonal malaria transmission in Burkina Faso, West Africa. *Infection and Immunity,* 76**,** 759-766.

NEWBOLD, C., CRAIG, A., KYES, S., ROWE, A., FERNANDEZ-REYES, D. & FAGAN, T. (1999) Cytoadherence, pathogenesis and the infected red cell surface in *Plasmodium falciparum*. *International Journal for Parasitology,* 29**,** 927-937.

NEWBOLD, C., WARN, P., BLACK, G., BERENDT, A., CRAIG, A., SNOW, B., MSOBO, M., PESHU, N. & MARSH, K. (1997) Receptor-specific adhesion and clinical disease in *Plasmodium falciparum*. *American Journal of Tropical Medicine and Hygiene,* 57**,** 389-398.

NIANG, M., XUE, Y. Y. & PREISER, P. R. (2009) The *Plasmodium falciparum* STEVOR multigene family mediates antigenic variation of the infected erythrocyte. *PLoS Pathogens,* 5.

NIELSEN, M. A., STAALSOE, T., KURTZHALS, J. A. L., GOKA, B. Q., DODOO, D., ALIFRANGIS, M., THEANDER, T. G., AKANMORI, B. D. & HVIID, L. (2002) *Plasmodium falciparum* variant surface antigen expression varies between isolates causing severe and nonsevere malaria and is modified by acquired immunity. *Journal of Immunology,* 168**,** 3444-3450.

NIKOLAIDIS, N. & NEI, M. (2004) Concerted and nonconcerted evolution of the Hsp70 gene superfamily in two sibling species of nematodes. *Molecular Biology and Evolution,* 21**,** 498-505.

NIMMERJAHN, F., BRUHNS, P., HORIUCHI, K. & RAVETCH, J. V. (2005) Fc gamma RIV: a novel FcR with distinct IgG subclass specificity. *Immunity,* 23**,** 41-51.

NIMMERJAHN, F. & RAVETCH, J. V. (2005) Divergent immunoglobulin g subclass activity through selective Fc receptor binding. *Science,* 310**,** 1510-1512.

NIMMERJAHN, F. & RAVETCH, J. V. (2006) Fc gamma receptors: old friends and new family members. *Immunity,* 24**,** 19-28.

NING, Z., COX, A., J. & MULLIKIN, J., C. (2001) SSAHA: a fast search method for large DNA databases. *Genome research***,** 1725-9.

NOSTEN, F., MCGREADY, R., SIMPSON, J. A., THWAI, K. L., BALKAN, S., CHO, T., HKIRIJAROEN, L., LOOAREESUWAN, S. & WHITE, N. J. (1999) Effects of *Plasmodium vivax* malaria in pregnancy. *Lancet,* 354**,** 546-549.

O'DEA, K. P., MCKEAN, P. G., JARRA, W. & BROWN, K. N. (1996) A single gene copy merozoite surface antigen and immune evasion? *Parasite Immunology,* 18**,** 165-172.

O'DONNELL, R. A., DE KONING-WARD, T. F., BURT, R. A., BOCKARIE, M., REEDER, J. C., COWMAN, A. F. & CRABB, B. S. (2001) Antibodies against merozoite surface protein (MSP)-1(19) are a major component of the invasion inhibitory response in individuals immune to malaria. *Journal of Experimental Medicine,* 193**,** 1403–1412.

OBI, R. K., OKANGBA, C. C., NWANEBU, F. C., NDUBUISI, U. U. & ORJI, N. M. (2011) Premunition in plasmodium falciparum malaria. *African Journal of Biotechnology,* 9**,** 1397-1401.

OCHOLA, L. B., SIDDONDO, B. R., OCHOLLA, H., NKYA, S., KIMANI, E. N., WILLIAMS, T. N., MAKALE, J. O., LILJANDER, A., URBAN, B. C., BULL, P. C., SZESTAK, T., MARSH, K. & CRAIG, A. G. (2011) Specific receptor usage in *Plasmodium falciparum* cytoadherence is associated with disease outcome. *PLoS One,* 6.

OCKENHOUSE, C. F., HO, M., TANDON, N. N., VAN SEVENTER, G. A., SHAW, S., WHITE, N. J., JAMIESON, G. A., CHULAY, J. D. & WEBSTER, H. K. (1991) Molecular Basis of Sequestration in Severe and Uncomplicated *Plasmodium falciparum* Malaria: Differential Adhesion of Infected Erythrocytes to CD36 and ICAM-1. *Journal of Infectious Diseases,* 164**,** 163-169.

OCKENHOUSE, C. F., TEGOSHI, T., MAENO, Y., BENJAMIN, C., HO, M., KHIN EI, K., THWAY, Y., WIN, K., AIKAWA, M. & LOBB, R. R. (1992) Human vascular endothelial cell adhesion receptors for *Plasmodium falciparum*-infected erythrocytes: Roles for endothelial leukocyte adhesion molecule 1 and vascular cell adhesion molecule 1. *Journal of Experimental Medicine,* 176**,** 1183-1189.

ODA, K., MATSUDA, H., MURAKAMI, T., KATAYAMA, S., OHGITANI, T. & YOSHIKAWA, M. (2000) Adjuvant and haemolytic activities of 47 saponins derived from medicinal and food plants. *Biological Chemistry,* 381**,** 67-74.

OFORI, M. F., DODOO, D., STAALSOE, T., KURTZHALS, J. A. L., KORAM, K., THEANDER, T. G., AKANMORI, B. D. & HVIID, L. (2002) Malaria-induced acquisition of antibodies to *Plasmodium falciparum* variant surface antigens. *Infection and Immunity,* 70**,** 2982-2988.

OGUTU, B. R., APOLLO, O. J., MCKINNEY, D., OKOTH, W., SIANGLA, J., DUBOVSKY, F., TUCKER, K., WAITUMBI, J. N., DIGGS, C., WITTES, J., MALKIN, E., LEACH, A., SOISSON, L. A., MILMAN, J. B., OTIENO, L., HOLLAND, C. A., POLHEMUS, M., REMICH, S. A., OCKENHOUSE, C. F., COHEN, J., BALLOU, W. R., MARTIN, S. K., ANGOV, E., STEWART, V. A., LYON, J. A., HEPPNER, D. G. & WITHERS, M. R. (2009) Blood stage malaria vaccine eliciting high antigen-specific antibody concentrations confers no protection to young children in Western Kenya. *PLoS One,* 4**,** e4708.

OKECH, B. A., CORRAN, P. H., TODD, J., JOYNSON-HICKS, A., UTHAIPIBULL, C., EGWANG, T. G., HOLDER, A. A. & RILEY, E. M. (2004) Fine specificity of serum antibodies to *Plasmodium falciparum* merozoite surface protein, PfMSP-1(19), predicts protection from malaria infection and high-density parasitemia. *Infection and Immunity,* 72**,** 1557-1567.

OLIVEIRA, T. R., FERNANDEZ-BECERRA, C., JIMENEZ, M. C. S., DEL PORTILLO, H. A. & SOARES, I. S. (2006) Evaluation of the acquired immune responses to *Plasmodium vivax* VIR variant antigens in individuals living in malaria-endemic areas of Brazil. *Malaria Journal,* 5.

OO, M. M., AIKAWA, M. & THAN, T. (1987) Human cerebral malaria: A pathological study. *Journal of Neuropathology and Experimental Neurology,* 46**,** 223-231.

ORTOLANO, F., MAFFIA, P., DEVER, G., HUTCHISON, S., BENSON, R. R., MILLINGTON, O. R., DE SIMONI, M. G., BUSHELL, T. J., GARSIDE, P., CARSWELL, H. V. & BREWER, J. M. (2009) Imaging T-cell movement in the brain during experimental cerebral malaria *Parasite Immunology,* 31**,** 147-150.

OSIER, F. H. A., WEEDALL, G. D., VERRA, F., MURUNGI, L., TETTEH, K. K. A., BULL, P., FABER, B. W., REMARQUE, E., THOMAS, A., MARSH, K. & CONWAY, D. J. (2010) Allelic Diversity and Naturally Acquired Allele-Specific Antibody Responses to *Plasmodium falciparum* Apical Membrane Antigen 1 in Kenya. *Infection and Immunity,* 78**,** 4625-4633.

OTTO, T. D., WILINSKI, D., ASSEFA, S., KEANE, T. M., SARRY, L. R., BÖHME, U., LEMIEUX, J., BARRELL, B., PAIN, A., BERRIMAN, M., NEWBOLD, C. & LLINÁS, M. (2010) New insights into the blood-stage transcriptome of *Plasmodium falciparum* using RNA-Seq. *Molecular Microbiology,* 76**,** 12-24.

OZSOLAK, F., GOREN, A., GYMREK, M., GUTTMAN, M., REGEV, A., BERNSTEIN, B. E. & MILOS, P. M. (2010) Digital transcriptome profiling from attomole-level RNA samples. *Genome Research,* 20**,** 519-25.

PAIN, A., BÖHME, U., BERRY, A. E., MUNGALL, K., FINN, R. D., JACKSON, A. P., MOURIER, T., MISTRY, J., PASINI, E. M., ASLETT, M. A., BALASUBRAMMANIAM, S., BORGWARDT, K., BROOKS, K., CARRET, C., CARVER, T. J., CHEREVACH, I., CHILLINGWORTH, T., CLARK, T. G., GALINSKI, M. R., HALL, N., HARPER, D., HARRIS, D., HAUSER, H., IVENS, A., JANSSEN, C. S., KEANE, T., LARKE, N., LAPP, S., MARTI, M., MOULE, S., MEYER, I. M., ORMOND, D., PETERS, N., SANDERS, M., SANDERS, S., SARGEANT, T. J., SIMMONDS, M., SMITH, F., SQUARES, R., THURSTON, S., TIVEY, A. R., WALKER, D., WHITE, B., ZUIDERWIJK, E., CHURCHER, C., QUAIL, M. A., COWMAN, A. F., TURNER, C. M. R., RAJANDREAM, M. A., KOCKEN, C. H. M., THOMAS, A. W., NEWBOLD, C. I., BARRELL, B. G. & BERRIMAN, M. (2008) The genome of the simian and human malaria parasite *Plasmodium knowlesi*. *Nature,* 455**,** 799-803.

PAIN, A., URBAN, B. C., KAI, O., CASALS-PASCUAL, C., SHAFI, J., MARSH, K. & ROBERTS, D. J. (2001) A non-sense mutation in CD36 gene is associated with protection from severe malaria. *Lancet,* 357**,** 1502-1503.

PAUL, F., ROATH, S. & MELVILLE, D. (1981) Separation of malaria-infected erythrocytes from whole blood: Use of a selective high-gradient magnetic separation technique. *Lancet,* 2**,** 70-71.

PAUL, W., E., BROWN, M., HORNBECK, P., MIZUGUCHI, J., OHARA, J., RABIN, E., SNAPPER, C. & TSANG, W. (1987) Regulation of B-lymphocyte activation, proliferation, and differentiation. . *Annals of the New York Academy of Sciences,* 505**,** 82-9.

PERRANT, R., MARRAMA, L., DIOUF, B., SOKHNA, C., TALL, A., NABETH, P., TRAPE, J. F., LONGACRE, S. & MERCEREAU-PUIJALON, O. (2005) Antibodies to the conserved C-terminal domain of the *Plasmodium falciparum* merozoite surface protein 1 and to the merozoite extract and their relationship with *in vitro* inhibitory

antibodies and protection against clinical malaria in a Senegalese village. *Journal of Infectious Diseases,* 191**,** 264-271.

PETES, T. D., MALONE, R. E. & SYMINGTON, L. S. (1991) Recombination in yeast. IN BROACH, J., JONES, E. & PRINGLE, J. (Eds.) *The molecular and cellular biology of the yeast Saccharomyces: genome dynamics, protein synthesis and energetics.* New York, Cold Spring Harbor Laboratory Press.

PETTER, M., BONOW, I. & KLINKERT, M. Q. (2008) Diverse expression patterns of subgroups of the *rif* multigene family during *Plasmodium falciparum* gametocytogenesis. *PLoS One,* 3**,** e3779.

PETTER, M., HAEGGSTRÖM, M., KHATTAB, A., FERNANDEZ, V., KLINKERT, M. Q. & WAHLGREN, M. (2007) Variant proteins of the *Plasmodium falciparum* RIFIN family show distinct subcellular localization and developmental expression patterns. *Molecular and Biochemical Parasitology,* 156**,** 51-61.

PFAFFL, M. W. (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Research,* 29.

PHILLIPS, R. E., LOOAREESUWAN, S. & WARRELL, D. A. (1986) The importance of anaemia in cerebral and uncomplicated *P. falciparum* malaria: Role of complications, dyserythropoiesis and iron sequestration. *Quarterly Journal of Medicine,* 58**,** 305-323.

PHILLIPS, R. S., BRANNAN, L. R., BALMER, P. & NEUVILLE, P. (1997) Antigenic variation during malaria infection - The contribution from the murine parasite Plasmodium chabaudi. *Parasite Immunology,* 19**,** 427-434.

PHILLIPS, R. S. & JONES, V. E. (1972) Immunity to *Plasmodium berghei* in rats: maximum levels of protective antibody activity are associated with eradication of the infection. *Parasitology,* 64**,** 117-127.

PHIRI, H., MONTGOMERY, J., MOLYNEUX, M. & CRAIG, A. G. (2009) Competitive endothelial adhesion between *Plasmodium falciparum* isolates under physiological flow conditions. *Malaria Journal,* 8**,** 214-223.

PINHEIRO, V. B. & ELLAR, D. J. (2006) How to kill a mocking bug? *Cellular Microbiology,* 8**,** 545-557.

PIZZI, E. & FRONTALI, C. (2001) Fine structure of *Plasmodium falciparum* subtelomeric sequences. *Molecular and Biochemical Parasitology,* 118**,** 253-258.

PODOBA, J. E. & STEVENSON, M. M. (1991) CD4+ and CD8+ T lymphocytes both contribute to acquired immunity to blood-stage *Plasmodium chabaudi* AS. *Infection and Immunity,* 59**,** 51-58.

PODZIMEK, S. (Ed.) (2011) *Light Scattering, in Light Scattering, Size Exclusion Chromatography and Asymmetric Flow Field Flow Fractionation: Powerful Tools for the Characterization of Polymers, Proteins and Nanoparticles,* Hoboken, New Jersey, USA, John Wiley & Sons, Inc.

POLLEY, S., D., MWANGI, T., KOCKEN, C., H., THOMAS, A. W., DUTTA, S., LANAR, D. E., REMARQUE, E., ROS, S. A., WILLIAMS, T. N., MWAMBINGU, G., LOWE, B., CONWAY, D. J. & MARSH, K. (2004) Human antibodies to recombinant protein constructs of *Plasmodium falciparum* Apical Membrane Antigen 1 (AMA1) and their associations with protection from malaria. *Vaccine,* 23**,** 718–728.

POUVELLE, B., GORMLEY, J. & TARASCHI, T. (1994) Characterization of trafficking pathways and membrane genesis in malaria-infected erythrocytes. *Molecular and Biochemical Parasitology,* 66**,** 83-96.

PRAKASH, D., FESEL, C., JAIN, R., CAZENAVE, P.-A., MISHRA, G. C. & PIED, S. (2006) Clusters of cytokines determine malaria severity in *Plasmodium falciparum*-infected patients from endemic areas of Central India. *Journal of Infectious Diseases,* 194**,** 198-207.

PRICE, R. N., DOUGLAS, N. M. & ANSTEY, N. M. (2009) New developments in *Plasmodium vivax* malaria: severe disease and the rise of chloroquine resistance. *Current Opinion in Infectious Diseases,* 22**,** 430-435.

PRUCCA, C. G., SLAVIN, I., QUIROGA, R., ELIÃÅAS, E. V., RIVERO, F. D., SAURA, A., CARRANZA, P. G. & LUJAÃÅN, H. D. (2008) Antigenic variation in *Giardia lamblia* is regulated by RNA interference. *Nature,* 456**,** 750-754.

PRZYBORSKI, J. M., MILLER, S. K., PFAHLER, J. M., HENRICH, P. P., ROHRBACH, P., CRABB, B. S. & LANZER, M. (2005) Trafficking of STEVOR to the Maurer's clefts in *Plasmodium falciparum*-infected erythrocytes. *EMBO Journal,* 24**,** 2306-2317.

PURKERSON, J. M. & ISAKSON, P. C. (1992) Interleukin 5 (IL-5) provides a signal that is required in addition to IL-4 for isotype switching to immunoglobulin (Ig) G1 and IgE. *Journal of Experimental Medicine,* 175**,** 973-82.

QUAIL, M. A., KOZAREWA, I., SMITH, F., SCALLY, A., STEPHENS, P. J., DURBIN, R., SWERDLOW, H. & TURNER, D. J. (2008) A large genome center's improvements to the Illumina sequencing system. *Nature Methods,* 5**,** 1005-1010.

QUIN, S. J. & LANGHORNE, J. (2001) Different regions of the malaria merozoite surface protein 1 of Plasmodium chabaudi elicit distinct T-cell and antibody isotype responses. *Infection and Immunity,* 69**,** 2245-2251.

QUINN, T. C. & WYLER, D. J. (1979) Mechanisms of action of hyperimmune serum in mediating protective immunity to rodent malaria (*Plasmodium berghei*). *Journal of Immunology,* 123**,** 2245-2249.

RALPH, S. A., SCHEIDIG-BENATAR, C. & SCHERF, A. (2005) Antigenic variation in *Plasmodium falciparum* is associated with movement of *var* loci between subnuclear locations. *Proceedings of the National Academy of Sciences of the United States of America,* 102**,** 5414-5419.

RASTI, N., WAHLGREN, M. & CHEN, Q. (2004) Molecular aspects of malaria pathogenesis. *FEMS Immunology and Medical Microbiology,* 41**,** 9-26.

RECKER, M., BUCKEE, C. O., SERAZIN, A., KYES, S., PINCHES, R., CHRISTODOULOU, Z., SPRINGER, A. L., GUPTA, S. & NEWBOLD, C. I. (2011) Antigenic Variation in *Plasmodium falciparum* Malaria Involves a Highly Structured Switching Pattern. *PLoS Pathogens,* 7**,** e1001306.

RECKER, M., NEE, S., BULL, P. C., KINYANJUI, S., MARSH, K., NEWBOLD, C. & GUPTA, S. (2004) Transient cross-reactive immune responses can orchestrate antigenic variation in malaria. *Nature,* 429**,** 555-8.

REEDER, J. C., ROGERSON, S. J., AL-YAMAN, F., ANDERS, R. F., COPPEL, R. L., NOVAKOVIC, S., ALPERS, M. P. & BROWN, G. V. (1994) Diversity of agglutinating phenotype, cytoadherence, and rosette-forming characteristics of *Plasmodium falciparum* isolates from Papua New Guinean children. *American Journal of Tropical Medicine and Hygiene,* 51**,** 45-55.

RILEY, E. M., ALLEN, S. J., WHEELER, J. G., J., B. M., BENNETT, S., TAKACS, B., SCHONFELD, H. J., A., H. A. & GREENWOOD, B. M. (1992) Naturally acquired cellular and humoral immune responses to the major merozoite surface antigen (*Pf* MSP1) of *Plasmodium falciparum* are associated with reduced malaria morbidity. *Parasite Immunology,* 14**,** 321–337.

RILEY, E. M., COUPER, K. N., HELMBY, H., HAFALLA, J. C. R., SOUZA, J. B. D., LANGHORNE, J., JARRA, W. & ZAVALA, F. (2010) Neuropathogenesis of human and murine malaria. *Trends in Parasitology,* 26**,** 277-278.

RIVERO, F. D., SAURA, A., PRUCCA, C. G., CARRANZA, P. G., TORRI, A. & LUJAN, H. D. (2010) Disruption of antigenic variation is crucial for effective parasite vaccine. *Nature Medicine,* 16**,** 551-557.

ROBERTS, D. J., CRAIG, A. G., BERENDT, A. R., PINCHES, R., NASH, G., MARSH, K. & NEWBOLD, C. I. (1992) Rapid switching to multiple antigenic and adhesive phenotypes in malaria. *Nature,* 357**,** 689-692.

ROBINSON, B. A., WELCH, T. L. & SMITH, J. D. (2003) Widespread functional specialization of *Plasmodium falciparum* erythrocyte membrane protein 1 family members to bind CD36 analysed across a parasite genome. *Molecular Microbiology,* 47**,** 1265-78.

ROES, J. & RAJEWSKY, K. (1991) Cell autonomous expression of IgD is not essential for the maturation of conventional B cells. *International Immunology,* 3**,** 1367-1371.

ROGERSON, S. J., HVIID, L., DUFFY, P. E., LEKE, R. F. & TAYLOR, D. W. (2007) Malaria in pregnancy: pathogenesis and immunity. *Lancet Infectious Diseases,* 7**,** 105-117.

ROGERSON, S. J., TEMBENU, R., DOBAÑO, C., PLITT, S., TAYLOR, T. E. & MOLYNEUX, M. E. (1999) Cytoadherence characteristics of *Plasmodium falciparum*-infected erythrocytes from Malawian children with severe and uncomplicated malaria. *American Journal of Tropical Medicine and Hygiene,* 61**,** 467-472.

ROTMAN, H. L., DALY, T. M., CLYNES, R. & LONG, C. A. (1998) Fc receptors are not required for antibody mediated protection against lethal malaria challenge in a mouse model. *Journal of Immunology,* 161**,** 1908-1912.

ROTTMANN, M., LAVSTSEN, T., MUGASA, J. P., KAESTLI, M., JENSEN, A. T. R., MÜLLER, D., THEANDER, T. & BECK, H. P. (2006) Differential expression of *var* gene groups is associated with morbidity caused by *Plasmodium falciparum* infection in Tanzanian children. *Infection and Immunity,* 74**,** 3904-3911.

ROWE, A., OBEIRO, J., NEWBOLD, C. I. & MARSH, K. (1995) *Plasmodium falciparum* rosetting is associated with malaria severity in Kenya. *Infection and Immunity,* 63**,** 2323-2326.

ROWE, A. J., MOULDS, J. M., NEWBOLD, C. I. & MILLER, L. H. (1997) *P. falciparum* rosetting mediated by a parasite-variant erythrocyte membrane protein and complement-receptor 1. *Nature,* 388**,** 292-295.

ROWE, A. K., ROWE, S. Y., SNOW, R. W., KORENROMP, E. L., ARMSTRONG SCHELLENBERG, J. R. M., STEIN, C., NAHLEN, B. L., BRYCE, J., BLACK, R. E. & STEKETEE, R. W. (2006) The burden of malaria mortality among African children in the year 2000. *International Journal of Epidemiology,* 35**,** 691-704.

ROWE, J. A., ROGERSON, S. J., RAZA, A., MOULDS, J. M., KAZATCHKINE, M. D., MARSH, K., NEWBOLD, C. I., ATKINSON, J. P. & MILLER, L. H. (2000) Mapping of the region of complement receptor (CR) 1 required for *Plasmodium falciparum* rosetting and demonstration of the importance of Cr1 in rosetting in field isolates. *Journal of Immunology,* 165**,** 6341-6346.

ROZEN, S. & SKALETSKY, H. (2000) Primer3 on the WWW for general users and for biologist programmers. IN CLIFTON, N. J. (Ed.) *Methods in molecular biology.* New Jersey, Humana Press.

RUDOLF, G. D. M. (1927) *Therapeutic malaria*.

RUSSELL, C., MERCEREAU-PUIJALON, O., LE SCANF, C., STEWARD, M. & ARNOT, D. E. (2005) Further definition of *Pf*EMP-1 DBL-1α domains mediating rosetting adhesion of *Plasmodium falciparum*. *Molecular and Biochemical Parasitology,* 144**,** 109-113.

RUTHERFORD, K., PARKHILL, J., CROOK, J., HORSNELL, T., RICE, P., RAJANDREAM, M. A. & BARRELL, B. (2000) Artemis: sequence visualization and annotation. *Bioinformatics,* 16**,** 944-5.

RUWENDE, C., KHOO, S. C., SNOW, R. W., YATES, S. N. R., KWIATKOWSKI, D., GUPTA, S., WARN, P., ALLSOPP, C. E. M., GILBERT, S. C., PESCHU, N., NEWBOLD, C. I., GREENWOOD, B. M., MARSH, K. & HILL, A. V. S. (1995) Natural selection of hemi- and heterozygotes for G6PD deficiency in Africa by resistance to severe malaria. *Nature,* 376**,** 246-249.

SABCHAREON, A., BURNOUF, T., OUATTARA, D., ATTANATH, P., BOUHAROUN-TAYOUN, H., CHANTAVANICH, P., FOUCAULT, C., CHONGSUPHAJAISIDDHI, T. & DRUILHE, P. (1991) Parasitologic and clinical human response to immunoglobulin administration in *falciparum* malaria. *American Journal of Tropical Medicine and Hygiene,* 45**,** 297-308.

SACHDEVA, G., KUMAR, K., JAIN, P. & RAMACHANDRAN, S. (2005) SPAAN: a software program for prediction of adhesins and adhesin-like proteins using neural networks. *Bioinformatics,* 21**,** 483-491.

SAFEUKUI, I., CORREAS, J. M., V., B., HIRT, D., DEPLAINE, G., MULÉ, S., LESURTEL, M., GOASGUEN, N., SAUVANET, A., COUVELARD, A., KERNEIS, S., KHUN, H., VIGAN-WOMAS, I., OTTONE, C., MOLINA, T. J., TRÉLUYER, J. M., MERCEREAU-PUIJALON, O., MILON, G., DAVID, P. H. & BUFFET, P. A. (2008) Retention of *Plasmodium falciparum* ring-infected erythrocytes in the slow, open microcirculation of the human spleen. *Blood,* 112**,** 2520-2528.

SAGARA, I., DICKO, A., ELLIS, R. D., FAY, M. P., DIAWARA, S. I., ASSADOU, M. H., SISSOKO, M. S., KONE, M., DIALLO, A. I., SAYE, R., GUINDO, M. A., KANTE, O., NIAMBELE, M. B., MIURA, K., MULLEN, G. E., PIERCE, M., MARTIN, L. B., DOLO, A., DIALLO, D. A., DOUMBO, O. K., MILLER, L. H. & SAUL, A. (2009) A randomized controlled phase 2 trial of the blood stage AMA1-C1/Alhydrogel malaria vaccine in children in Mali. *Vaccine,* 27**,** 3090-3098.

SAKIHAMA, N., KIMURA, M., HIRAYAMA, K., KANDA, T., NA-BANGCHANG, K., JONGWUTIWES, S., CONWAY, D. & TANABE, K. (1999) Allelic recombination and linkage disequilibrium within MSP-1 of *Plasmodium falciparum*, the malignant human malaria parasite. *Gene,* 230**,** 47-54.

SALANTI, A., DAHLBÄCK, M., TURNER, L., NIELSEN, M. A., BARFOD, L., MAGISTRADO, P., JENSEN, A. T. R., LAVSTSEN, T., OFORI, M. F., MARSH, K., HVIID, L. & THEANDER, T. G. (2004) Evidence for the involvement of VAR2CSA in pregnancy-associated malaria. *Journal of Experimental Medicine,* 200**,** 1197-1203.

SAM, H., SU, Z. & STEVENSON, M. M. (1999) Deficiency in tumor necrosis factor alpha activity does not impair early protective Th1 responses against blood-stage malaria. *Infection and Immunity,* 67**,** 2660-2664.

SAM-YELLOWE, T. Y. (2009) The role of the Maurer's clefts in protein transport in *Plasmodium falciparum*. *Trends in Parasitology,* 25**,** 277-284.

SAM-YELLOWE, T. Y., FLORENCE, L., JOHNSON, J. R., WANG, T., DRAZBA, J. A., LE ROCH, K. G., ZHOU, Y., BATALOV, S., CARUCCI, D. J., WINZELER, E. A. & YATES III, J. R. (2004) A *Plasmodium* gene family encoding maurer's cleft membrane proteins: Structural properties and expression profiling. *Genome Research,* 14**,** 1052-1059.

SAMBROOK, J. & RUSSELL, D. W. (2001) *Molecular cloning : a laboratory manual,* New York, Cold Spring Harbor Laboratory.

SANNI, L., ALLSOPP, C., REUBSAET, L., SANNI, A., NEWBOLD, C., CHAUHAN, V. & LANGHORNE, J. (2002) Cellular responses to *Plasmodium falciparum* erythrocyte membrane protein-1: use of relatively conserved synthetic peptide pools to determine CD4 T cell responses in malaria-exposed individuals in Benin, West Africa. *Malaria Journal,* 1**,** 7.

SARGEANT, T. J., MARTI, M., CALER, E., CARLTON, J. M., SIMPSON, K., SPEED, T. P. & COWMAN, A. F. (2006) Lineage-specific expansion of proteins exported to erythrocytes in malaria parasites. *Genome Biology,* 7.

SARIDAKI, T., FRÖHLICH, K. S., BRAUN-BRETON, C. & LANZER, M. (2009) Export of *Pf*SBP1 to the *Plasmodium falciparum* Maurer's Clefts. *Traffic,* 10**,** 137-152.

SCHERF, A., FIGUEIREDO, L. & H., F.-J. L. (2004) Chromosome structure and dynamics of *Plasmodium* subtelomeres. IN WATERS, A. P. & JANSE, C. J. (Eds.) *Malaria Parasites: Genomes and Molecular Biology.* Caister Academic Press.

SCHERF, A., HERNANDEZ-RIVAS, R., BUFFET, P., BOTTIUS, E., BENATAR, C., POUVELLE, B., GYSIN, J. & LANZER, M. (1998) Antigenic variation in malaria: in situ switching, relaxed and mutually exclusive transcription of *var* genes during intra-erythrocytic development in *Plasmodium falciparum*. *EMBO Journal,* 17**,** 5418-26.

SCHERF, A., LOPEZ-RUBIO, J. J. & RIVIERE, L. (2008) Antigenic variation in *Plasmodium falciparum*. *Annual Review of Microbiology.*

SCHOBERT, B. & TSCHESCHE, H. (1978) Unusual solution properties of proline and its interaction with proteins. *Biochimica et Biophysica Acta,* 541**,** 270-277.

SCHREIBER, N., BRATTIG, N., EVANS, J., TSIRI, A., HORSTMANN, R. D., MAY, J. & KLINKERT, M. Q. (2006) Cerebral malaria is associated with IgG2 and IgG4

antibody responses to recombinant *Plasmodium falciparum* RIFIN antigen. *Microbes and Infection,* 8**,** 1269-76.

SCHREIBER, N., KHATTAB, A., PETTER, M., MARKS, F., ADJEI, S., KOBBE, R., MAY, J. & KLINKERT, M. Q. (2008) Expression of *Plasmodium falciparum* 3D7 STEVOR proteins for evaluation of antibody responses following malaria infections in naive infants. *Parasitology,* 135**,** 155-167.

SERGENT, E. & PARROT, L. (1935) L'immunité, la prémunition et la résistance innée. *Archives de l'Institut Pasteur d'Algérie Institut Pasteur d'Algérie,* 13**,** 279-319.

SEVIER, C. S. & KAISER, C. A. (2002) Formation and transfer of disulphide bonds in living cells. *Nature Reviews Molecular and Cellular Biology,* 3**,** 836-847.

SEXTON, A. C., GOOD, R. T., HANSEN, D. S., D'OMBRAIN, M. C., BUCKINGHAM, L., SIMPSON, K. & SCHOFIELD, L. (2004) Transcriptional profiling reveals suppressed erythropoiesis, up-regulated glycolysis, and interferon-associated responses in murine malaria. *Journal of Infectious Diseases,* 189**,** 1245-1256.

SHAKHAR, G., LINDQUIST, R. L., SKOKOS, D., DUDZIAK, D., HUANG, J. H., NUSSENZWEIG, M. C. & DUSTIN, M. L. (2005) Stable T cell - dendritic cell interactions precede the development of both tolerance and immunity *in vivo*. *Nature Immunology,* 6**,** 707-714.

SHARP, S., LAVSTSEN, T., FIVELMAN, Q. L., SAEED, M., MCROBERT, L., TEMPLETON, T. J., JENSEN, A. T. R., BAKER, D. A., THEANDER, T. G. & SUTHERLAND, C. J. (2006) Programmed transcription of the *var* gene family, but not of *stevor*, in *Plasmodium falciparum* gametocytes. *Eukaryotic Cell,* 5**,** 1206-1214.

SHENDURE, J. (2008) The beginning of the end for microarrays? *Nature Methods,* 5**,** 585-7.

SHI, Q., CERNETICH, A., DALY, T. M., GALVAN, G., VAIDYA, A. B., BERGMAN, L. W. & BURNS JR, J. M. (2005) Alteration in host cell tropism limits the efficacy of immunization with a surface protein of malaria merozoites. *Infection and Immunity,* 73**,** 6363-6371.

SHIMODAIRA, H. & HASEGAWA, M. (1999) Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Molecular Biology and Evolution,* 16**,** 1114–1116.

SHINKAIA, Y., RATHBUNA, G., LAMA, K. P., OLTZA, E. M., STEWARTA, V., MENDELSOHNA, M., CHARRONA, J., DATTAA, M., YOUNGA, F., STALLA, A. M. & ALTA, F. W. (1992) RAG-2-deficient mice lack mature lymphocytes owing to inability to initiate V(D)J rearrangement *Cell,* 68**,** 855-67.

SHIRLEY, B. A. (1995) *Protein stability and folding: theory and practice,* New Jersey, Humana Press.

SIDDIQUI, W. A., KAN, S. C., KRAMER, K. & RICHMOND-CRUM, S. M. (1979) Recent developments in production and purification of malaria antigens: *In vitro* production and partial purification of *Plasmodium falciparum* antigen. *WHO bulletin,* 57**,** 75 - 82.

SIJWALI, P. S. & ROSENTHAL, P. J. (2010) Functional evaluation of *Plasmodium* export signals in *Plasmodium berghei* suggests multiple modes of protein export. *PLoS One,* 5.

SILVERMAN, P. H., SCHOOLEY, J. C. & MAHLMANN, L. J. (1987) Murine malaria decreases hemopoietic stem cells. *Blood,* 69**,** 408-413.

SLADE, S. J. & LANGHORNE, J. (1989) Production of interferon-gamma during infection of mice with *Plasmodium chabaudi chabaudi*. *Immunobiology,* 179**,** 353-365.

SMITH, J. D., CRAIG, A. G., KRIEK, N., HUDSON-TAYLOR, D., KYES, S., FAGEN, T., PINCHES, R., BARUCH, D. I., NEWBOLD, C. I. & MILLER, L. H. (2000) Identification of a *Plasmodium falciparum* intercellular adhesion molecule-1 binding domain: A parasite adhesion trait implicated in cerebral malaria. *Proceedings of the National Academy of Sciences of the United States of America,* 97**,** 1766-1771.

SMYTHE, J. A., COPPEL, R. L., BROWN, G. V., RAMASAMY, R., KEMP, D. J. & ANDERS, R. F. (1988) Identification of two integral membrane proteins of

Plasmodium falciparum. *Proceedings of the National Academy of Sciences of the United States of America,* 85**,** 5195-9.

SNOUNOU, G., JARRA, W., VIRIYAKOSOL, S., WOOD, J. C. & BROWN, K. N. (1989) Use of a DNA probe to analyse the dynamics of infection with rodent malaria parasites confirms that parasite clearance during crisis is predominantly strain- and species-specific. *Molecular and Biochemical Parasitology***,** 37-46.

SONNHAMMER, E. L., EDDY, S. R., BIRNEY, E., BATEMAN, A. & DURBIN, R. (1998) Pfam: multiple sequence alignments and HMM-profiles of protein domains. *Nucleic Acids Research,* 26**,** 320–322.

SPACCAPELO, R., JANSE, C. J., CATERBI, S., FRANKE-FAYARD, B., BONILLA, J. A., SYPHARD, L. M., DI CRISTINA, M., DOTTORINI, T., SAVARINO, A., CASSONE, A., BISTONI, F., WATERS, A. P., DAME, J. B. & CRISANTI, A. (2010) Plasmepsin 4-deficient *Plasmodium berghei* are virulence attenuated and induce protective immunity against experimental malaria. *American Journal of Pathology,* 176**,** 205-217.

SPENCE, P. J., CUNNINGHAM, D., JARRA, W., LAWTON, J., LANGHORNE, J. & THOMPSON, J. (2011) Transformation of the rodent malaria parasite *Plasmodium chabaudi*. *Nature Protocols,* 6**,** 553-561.

SPIELMANN, T. & GILBERGER, T. W. (2010) Protein export in malaria parasites: do multiple export motifs add up to multiple export pathways? *Trends in Parasitology,* 26**,** 6-10.

SPONAAS, A. M., DO ROSARIO, A. P. F., VOISINE, C., MASTELIC, B., THOMPSON, J., KOERNIG, S., JARRA, W., RENIA, L., MAUDUIT, M., POTOCNIK, A. J. & LANGHORNE, J. (2009) Migrating monocytes recruited to the spleen play an important role in control of blood stage malaria. *Blood,* 114**,** 5522-5531.

SPRINGER, A. L., SMITH, L. M., MACKAY, D. Q., NELSON, S. O. & SMITH, J. D. (2004) Functional interdependence of the DBLbeta domain and c2 region for binding of the *Plasmodium falciparum* variant antigen to ICAM-1. *Molecular and Biochemical Parasitology,* 137**,** 55-64.

SPYCHER, C., KLONIS, N., SPIELMANN, T., KUMP, E., STEIGER, S., TILLEY, L. & BECK, H. P. (2003) MAHRP-1, a novel *Plasmodium falciparum* histidine-rich protein, binds ferriprotoporphyrin IX and localizes to the Maurer's clefts. *Journal of Biological Chemistry,* 278**,** 35373-35383.

SPYCHER, C., RUG, M., KLONIS, N., FERGUSON, D. J. P., COWMAN, A. F., BECK, H. P. & TILLEY, L. (2006) Genesis of and trafficking to the Maurer's clefts of *Plasmodium falciparum*-infected erythrocytes. *Molecular and Cellular Biology,* 26**,** 4074-4085.

SREERAMA, N. & WOODY, R. W. (2000) Estimation of protein secondary structure from circular dichroism spectra: Comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Analytical Biochemistry,* 287**,** 252-260.

SRIPRAWAT, K., KAEWPONGSRI, S., SUWANARUSK, R., LEIMANIS, M. L., LEK-UTHAI, U., PHYO, A. P., SNOUNOU, G., RUSSELL, B., RENIA, L. & NOSTEN, F. (2009) Effective and cheap removal of leukocytes and platelets from *Plasmodium vivax* infected blood. *Malaria Journal,* 8.

SRIVASTAVA, A., GANGNARD, S., ROUND, A., DECHAVANNE, S., JUILLERAT, A., RAYNAL, B., FAURE, G., BARON, B., RAMBOARINA, S., SINGH, S. K., BELRHALI, H., ENGLAND, P., LEWIT-BENTLEY, A., SCHERF, A., BENTLEY, G. A. & GAMAIN, B. (2010) Full-length extracellular region of the var2CSA variant of *Pf*EMP1 is required for specific, high-affinity binding to CSA. *Proceedings of the National Academy of Sciences of the United States of America,* 107**,** 4884-9.

STAALSOE, T., HAMAD, A. A., HVIID, L., ELHASSAN, I. M., ARNOT, D. E. & THEANDER, T. G. (2002) *In vivo* switching between variant surface antigens in human *Plasmodium falciparum* infection. *Journal of Infectious Diseases,* 186**,** 719-22.

STANISIC, D. I., RICHARDS, J. S., MCCALLUM, F. J., MICHON, P., KING, C. L., SCHOEPFLIN, S., GILSON, P. R., MURPHY, V. J., ANDERS, R. F., MUELLER, I.

& BEESON, J. G. (2009) Immunoglobulin G Subclass-Specific Responses against *Plasmodium falciparum* Merozoite Antigens Are Associated with Control of Parasitemia and Protection from Symptomatic Illness. *Infection and Immunity,* 77**,** 1165-1174.

STEPHENS, R., ALBANO, F. R., QUIN, S., PASCAL, B. J., HARRISON, V., STOCKINGER, B., KIOUSSIS, D., WELTZIEN, H. U. & LANGHORNE, J. (2005) Malaria-specific transgenic CD4+ T cells protect immunodeficient mice from lethal infection and demonstrate requirement for a protective threshold of antibody production for parasite clearance. *Blood,* 106**,** 1676-1684.

STEPHENS, R. & LANGHORNE, J. (2010) Effector memory Th1 CD4 T cells are maintained in a mouse model of chronic malaria. *PLoS Pathogens,* 6.

STEVENSON, M. M., GROS, P., OLIVIER, M., FORTIN, A. & SERGHIDES, L. (2010) Cerebral malaria: human versus mouse studies. *Trends in Parasitology,* 26**,** 274-275.

STEVENSON, M. M. & RILEY, E. M. (2004) Innate immunity to malaria. *Nature Reviews Immunology,* 4**,** 168-180.

STOUTE, J. A., ODINDO, A. O., OWUOR, B. O., MIBEI, E. K., OPOLLO, M. O. & WAITUMBI, J. N. (2003) Loss of red blood cell-complement regulatory proteins and increased levels of circulating immune complexes are associated with severe malarial anemia. *Journal of Infectious Diseases,* 187**,** 522-525.

SU, X.-Z. & WELLEMS, T. E. (1996) Towards a high resolution *Plasmodium falciparum* linkage map: Polymorphic markers from hundreds of simple sequence repeats. . *Genomics,* 33**,** 430-444.

SU, X. Z., FERDIG, M. T., HUANG, Y., HUYNH, C. Q., LIU, A., YOU, J., WOOTTON, J. C. & WELLEMS, T. E. (1999) A genetic map and recombination parameters of the human malaria parasite *Plasmodium falciparum*. *Science,* 286**,** 1351-1353.

SU, Z., FORTIN, A., GROS, P. & STEVENSON, M. M. (2002) Opsonin-independent phagocytosis: An effector mechanism against acute blood-stage *Plasmodium chabaudi* AS infection. *Journal of Infectious Diseases,* 186**,** 1321-1329.

TATEM, A., J. , SMITH, D., L., GETHING, P., W., KABARIA, C., W., SNOW, R. W. & HAY, S., I. (2010) Ranking of elimination feasibility between malaria-endemic countries. *Lancet,* 376**,** 1579-1591.

TAYLOR, H. M., KYES, S. A., HARRIS, D., KRIEK, N. & NEWBOLD, C. I. (2000a) A study of *var* gene transcription *in vitro* using universal *var* gene primers. *Molecular and Biochemical Parasitology,* 105**,** 13-23.

TAYLOR, H. M., KYES, S. A. & NEWBOLD, C. I. (2000b) *var* gene diversity in *Plasmodium falciparum* is generated by frequent recombination events. *Molecular and Biochemical Parasitology,* 110**,** 391-397.

TAYLOR, P. R., SEIXAS, E., WALPORT, M. J., LANGHORNE, J. & BOTTO, M. (2001) Complement contributes to protective immunity against reinfection by *Plasmodium chabaudi chabaudi* parasites. *Infection and Immunity,* 69**,** 3853-3859.

TAYLOR, T. E., BORGSTEIN, A. & MOLYNEUX, M. E. (1993) Acid-base status in paediatric *Plasmodium falciparum* malaria. *Quarterly Journal of Medicine,* 86**,** 99-109.

TAYLOR, T. E., FU, W. J., CARR, R. A., WHITTEN, R. O., MUELLER, J. G., FOSIKO, N. G., LEWALLEN, S., LIOMBA, N. G. & MOLYNEUX, M. E. (2004) Differentiating the pathologies of cerebral malaria by postmortem parasite counts. *Nature Medicine,* 10**,** 143-145.

THAM, W. H., PAYNE, P. D., BROWN, G. V. & ROGERSON, S. J. (2007) Identification of basic transcriptional elements required for *rif* gene transcription. *International Journal for Parasitology,* 37**,** 605-615.

TIAN, J., KUMAR, S., KASLOW, D. & MILLER, L. (1997) Comparison of protection induced by immunization with recombinant proteins from different regions of merozoite surface protein 1 of *Plasmodium yoelii*. *Infection and Immunity,* 65**,** 3032-3036.

TRAPE, J. F. (2001) The public health impact of chloroquine resistance in Africa. *American Journal of Tropical Medicine and Hygiene,* 64**,** 12-17.

TRAPNELL, C., PACHTER, L. & SALZBERG, S. L. (2009) TopHat: Discovering splice junctions with RNA-Seq. *Bioinformatics,* 25**,** 1105-1111.

TREECK, M., ZACHERL, S., HERRMANN, S., CABRERA, A., KONO, M., STRUCK, N. S., ENGELBERG, K., HAASE, S., FRISCHKNECHT, F., MIURA, K., SPIELMANN, T. & GILBERGER, T. W. (2009) Functional analysis of the leading malaria vaccine candidate AMA-1 reveals an essential role for the cytoplasmic domain in the invasion process. *PLoS Pathogens,* 5.

TSCHOPP, J. F., BRUST, P. F., CREGG, J. M., STILLMAN, C. & GINGERAS, T. R. (1987) Expression of the lacZ Gene from Two Methanol Regulated Promoters in *Pichia pastoris*. *Nucleic Acids Research,* 15**,** 3859-3876.

TSUBOI, T., TAKEO, S., IRIKO, H., JIN, L., TSUCHIMOCHI, M., MATSUDA, S., HAN, E., T., OTSUKI, H., KANEKO, O., SATTABONGKOT, J., UDOMSANGPETCH, R., SAWASAKI, T., TORII, M. & ENDO, Y. (2008) Wheat germ cell-free system-based production of malaria proteins for discovery of novel vaccine candidates. *Infection and Immunity,* 76**,** 1702-8.

TURNER, C. M. R. (2002) A perspective on clonal phenotypic (antigenic) variation in protozoan parasites. *Parasitology,* 125**,** S17-S23.

TURNER, G. D. H., MORRISON, H., JONES, M., DAVIS, T. M. E., LOOAREESUWAN, S., BULEY, I. D., GATTER, K. C., NEWBOLD, C. I., PUKRITAYAKAMEE, S., NAGACHINTA, B., WHITE, N. J. & BERENDT, A. R. (1994) An immunohistochemical study of the pathology of fatal malaria: Evidence for widespread endothelial activation and a potential role for intercellular adhesion molecule-1 in cerebral sequestration. *American Journal of Pathology,* 145**,** 1057-1069.

UTHAIPIBULL, C., AUFIERO, B., SYED, S. E. H., HANSEN, B., PATIÑO, J. A. G., ANGOV, E., LING, I. T., FEGEDING, K., MORGAN, W. D., OCKENHOUSE, C., BIRDSALL, B., FEENEY, J., LYON, J. A. & HOLDER, A. A. (2001) Inhibitory and blocking monoclonal antibody epitopes on merozoite surface protein 1 of the malaria parasite *Plasmodium falciparum*. *Journal of Molecular Biology,* 307**,** 1381-1394.

VAN DER HEYDE, H. C., NOLAN, J., COMBES, V., GRAMAGLIA, I. & GRAU, G. E. (2006) A unified hypothesis for the genesis of cerebral malaria: sequestration, inflammation and hemostasis leading to microcirculatory dysfunction. *Trends in Parasitology,* 22**,** 503-508.

VAN NOORT, S. P., NUNES, M. C., WEEDALL, G. D., HVIID, L. & GOMES, M. G. (2010) Immune selection and within-host competition can structure the repertoire of variant surface antigens in *Plasmodium falciparum*-a mathematical model. *PLoS One,* 5**,** e9778.

VAN OOIJ, C., TAMEZ, P., BHATTACHARJEE, S., HILLER, N. L., HARRISON, T., LIOLIOS, K., KOOIJ, T., RAMESAR, J., BALU, B., ADAMS, J., WATERS, A., JANSE, C. & HALDAR, K. (2008) The malaria secretome: From algorithms to essential function in blood stage infection. *PLoS Pathogens,* 4.

VELDHOEN, M., UYTTENHOVE, C., VAN SNICK, J., HELMBY, H., WESTENDORF, A., BUER, J., MARTIN, B., WILHELM, C. & STOCKINGER, B. (2008) Transforming growth factor-beta 'reprograms' the differentiation of T helper 2 cells and promotes an interleukin 9-producing subset. *Nature Immunology,* 9**,** 1341-1346.

VESTERGAARD, L. S., LUSINGU, J. P., NIELSEN, M. A., MMBANDO, B. P., DODOO, D., AKANMORI, B. D., ALIFRANGIS, M., BYGBJERG, I. C., LEMNGE, M. M., STAALSOE, T., HVIID, L. & THEANDER, T. G. (2008) Differences in human antibody reactivity to *Plasmodium falciparum* variant surface antigens are dependent on age and malaria transmission intensity in northeastern Tanzania. *Infection and Immunity,* 76**,** 2706-14.

VIGNALI, M., ARMOUR, C. D., CHEN, J., MORRISON, R., CASTLE, J. C., BIERY, M. C., BOUZEK, H., MOON, W., BABAK, T., FRIED, M., RAYMOND, C. K. & DUFFY, P. E. (2011) NSR-seq transcriptional profiling enables identification of a gene signature of *Plasmodium falciparum* parasites infecting children. *Journal of Clinical Investigation,* 121**,** 1119-1129.

VILLEVAL, J. L., LEW, A. & METCALF, D. (1990) Changes in hemopoietic and regulator levels in mice during fatal or nonfatalmalarial infections. I. Erythropoietic populations. *Experimental Parasitology,* 71**,** 364-374.

VILMOS, P. & KURUCZ, E. (1998) Insect immunity: Evolutionary roots of the mammalian innate immune system. *Immunology Letters,* 62**,** 59-66.

VINAROV, D. A., NEWMAN, C. L. L. & MARKLEY, J. L. (2006) Wheat germ cell-free platform for eukaryotic protein production. *FEBS Journal,* 273**,** 4160-4169.

VOSS, T. S., KAESTLI, M., VOGEL, D., BOPP, S. & BECK, H. P. (2003) Identification of nuclear proteins that interact differentially with *Plasmodium falciparum var* gene promoters. *Molecular Microbiology,* 48**,** 1593-1607.

VUKOVIC, P., HOGARTH, P., M , BARNES, N., KASLOW, D., C. & GOOD, M., F. (2000) Immunoglobulin G3 antibodies specific for the 19-kilodalton carboxyl-terminal fragment of *Plasmodium yoelii* merozoite surface protein 1 transfer protection to mice deficient in Fc-gammaRI receptors. *Infection and Immunity,* 68**,** 3019-22.

WALLER, K. L., COOKE, B. M., NUNOMURA, W., MOHANDAS, N. & COPPEL, R. L. (1999) Mapping the binding domains involved in the interaction between the *Plasmodium falciparum* knob-associated histidine-rich protein (KAHRP) and the cytoadherence ligand *P. falciparum* erythrocyte membrane protein 1 (*Pf*EMP1). *Journal of Biological Chemistry,* 274**,** 23808-23813.

WALTER, P. R., GARIN, Y. & BLOT, P. (1982) Placental pathologic changes in malaria. A histologic and ultrastructural study. *American Journal of Pathology,* 109**,** 330-342.

WANG, C. W., MWAKALINGA, S. B., SUTHERLAND, C. J., SCHWANK, S., SHARP, S., HERMSEN, C. C., SAUERWEIN, R. W., THEANDER, T. G. & LAVSTSEN, T. (2010) Identification of a major *rif* transcript common to gametocytes and sporozoites of *Plasmodium falciparum*. *Malaria Journal,* 9**,** 147.

WARD, G., E., MILLER, L., H. & DVORAK, J., A. (1993) The origin of parasitophorous vacuole membrane lipids in malaria infected erythrocytes. *Journal of Cell Science* 106**,** 237-248.

WARGO, A., R., RANDLE, N., CHAN, B., H., K., THOMPSON, J., READ, A., F. & BABIKER, H., A. (2006) *Plasmodium chabaudi*: Reverse transcription PCR for the detection and quantification of transmission stage malaria parasites. *Experimental Parasitology,* 112**,** 13-20.

WARIMWE, G. M., KEANE, T. M., FEGAN, G., MUSYOKI, J. N., NEWTON, C. R. J. C., PAIN, A., BERRIMAN, M., MARSH, K. & BULL, P. C. (2009) *Plasmodium falciparum var* gene expression is modified by host immunity. *Proceedings of the National Academy of Sciences of the United States of America,* 106**,** 21801-21806.

WASSMER, S., C., MOXON, C., A., TAYLOR, T. E., GRAU, G., E., MOLYNEUX, M., E. & CRAIG, A., G. (2011) Vascular endothelial cells cultured from patients with cerebral or uncomplicated malaria exhibit differential reactivity to TNF. *Cellular Microbiology,* 13**,** 198-209.

WASSMER, S. C., LÉPOLARD, C., TRAORÉ, B., POUVELLE, B., GYSIN, J. & GRAU, G. E. (2004) Platelets Reorient *Plasmodium falciparum*-Infected Erythrocyte Cytoadhesion to Activated Endothelial Cells. *Journal of Infectious Diseases,* 189**,** 180-189.

WEBER, J. L. (1988) Interspersed repetitive DNA from *Plasmodium falciparum*. *Molecular and Biochemical Parasitology,* 29**,** 117-124.

WEILL, J., C. & REYNAUD, C., A. (1996) Rearrangement/hypermutation/gene conversion: when, where and why? *Immunology Today,* 17**,** 92-97.

WEIR, B., S. & COCKERHAM, C., C. (1984) Estimating F statistics for the analysis of population structure. . *Evolution,* 38**,** 1358-1370.

WEISS, L., JOHNSON, J. & WEIDANZ, W. (1989) Mechanisms of splenic control of murine malaria: tissue culture studies of the erythropoietic interplay of spleen, bone marrow, and blood in lethal (strain 17XL) *Plasmodium yoelii* malaria in BALB/c mice. *American Journal of Tropical Medicine and Hygiene,* 41**,** 135-143.

WESTENBERGER, S. J., MCCLEAN, C. M., CHATTOPADHYAY, R., DHARIA, N. V., CARLTON, J. M., BARNWELL, J. W., COLLINS, W. E., HOFFMAN, S. L., ZHOU, Y., VINETZ, J. M. & WINZELER, E. A. (2010) A systems-based analysis of *Plasmodium vivax* lifecycle transcription from human to mosquito. *PLoS Neglected Tropical Diseases,* 4.

WHITE, N., J. (2008) *Plasmodium knowlesi*: The Fifth Human Malaria Parasite *Clinical Infectious Diseases,* 46**,** 172-173.

WHITE, W. I., EVANS, C. B. & TAYLOR, D. W. (1991) Antimalarial antibodies of the immunoglobulin G2a isotype modulate parasitemias in mice infected with *Plasmodium yoelii*. *Infection and Immunity,* 59**,** 3547-3554.

WHO (2008) *World Malaria Report.*

WICKHAM, M. E., RUG, M., RALPH, S. A., KLONIS, N., MCFADDEN, G. I., TILLEY, L. & COWMAN, A. F. (2001) Trafficking and assembly of the cytoadherence complex in *Plasmodium falciparum*-infected human erythrocytes. *EMBO Journal,* 20**,** 5636-5649.

WIESNER, J., JOMAA, H., WILHELM, M., TONY, H., P. , KREMSNER, P., G. , HORROCKS, P. & LANZER, M. (1997) Host cell factor CD59 restricts complement lysis of *Plasmodium falciparum*-infected erythrocytes. *European Journal of Immunology,* 27**,** 2708-2713.

WILLIAMS, C. T. & AZAD, A. F. (2010) Transcriptional analysis of the pre-erythrocytic stages of the rodent malaria parasite, *Plasmodium yoelii*. *PLoS One,* 5.

WINTER, G., KAWAI, S., HAEGGSTRÖM, M., KANEKO, O., VON EULER, A., KAWAZU, S. I., PALM, D., FERNANDEZ, V. & WAHLGREN, M. (2005) SURFIN is a polymorphic antigen expressed on *Plasmodium falciparum* merozoites and infected erythrocytes. *Journal of Experimental Medicine,* 201**,** 1853-1863.

XIAO, L., PATTERSON, P. S., YANG, C. & LAL, L. (1999) Role of eicosanoids in the pathogenesis of murine cerebral malaria. *American Journal of Tropical Medicine and Hygiene,* 60**,** 668-673.

XIAO, Y. P., AL-KHEDERY, B. & ALLRED, D. R. (2010) The *Babesia bovis* VESA1 virulence factor subunit 1b is encoded by the 1beta branch of the *ves* multigene family. *Molecular and Biochemical Parasitology,* 171**,** 81-88.

YADAVA, A. & OCKENHOUSE, C. F. (2003) Effect of codon optimization on expression levels of a functionally folded malaria vaccine candidate in prokaryotic and eukaryotic expression systems. *Infection and Immunity,* 71**,** 4961-4969.

YAMAUCHI, L., M, L. M., COPPI, A., SNOUNOU, G. & SINNIS, P. (2007) *Plasmodium* sporozoites trickle out of the injection site. *Cellular Microbiology,* 9**,** 1215-1222.

YEUNG, S., SOCHEAT, D., MOORTHY, V., S. & MILLS, A., J. (2009) Artemisinin resistance on the Thai-Cambodian border. *Lancet,* 374**,** 1418-1419.

YONE, C. L. R. P., KREMSNER, P. G. & LUTY, A. J. F. (2005) Immunoglobulin G isotype responses to erythrocyte surface-expressed variant antigens of *Plasmodium falciparum* predict protection from malaria in African children. *Infection and Immunity,* 73**,** 2281-2287.

YUDA, M., IWANAGA, S., SHIGENOBU, S., MAIR, G. R., JANSE, C. J., WATERS, A. P., KATO, T. & KANEKO, I. (2009) Identification of a transcription factor in the mosquito-invasive stage of malaria parasites. *Molecular Microbiology,* 71**,** 1402-1414.

ZEIG, J., SILVERMAN, M., HILMEN, M. & SIMON, M. (1977) Recombinational switch for gene expression. *Science,* 196**,** 170-172.

ZOU, L., MILES, A. P., WANG, J. & STOWERS, A. W. (2003) Expression of malaria transmission-blocking vaccine antigen Pfs25 in *Pichia pastoris* for use in human clinical trials. *Vaccine,* 21**,** 1650-1657.

ZUCKERKANDL, E. & PAULING, L. (1965) Molecules as documents of evolutionary history. *Journal of Theoretical Biology,* 8**,** 357-66.

ZUCKERMAN, A. (1957) Blood loss and replacement in plasmodial infection. *Journal of Infectious Diseases,* 100**,** 172-206.

Appendices

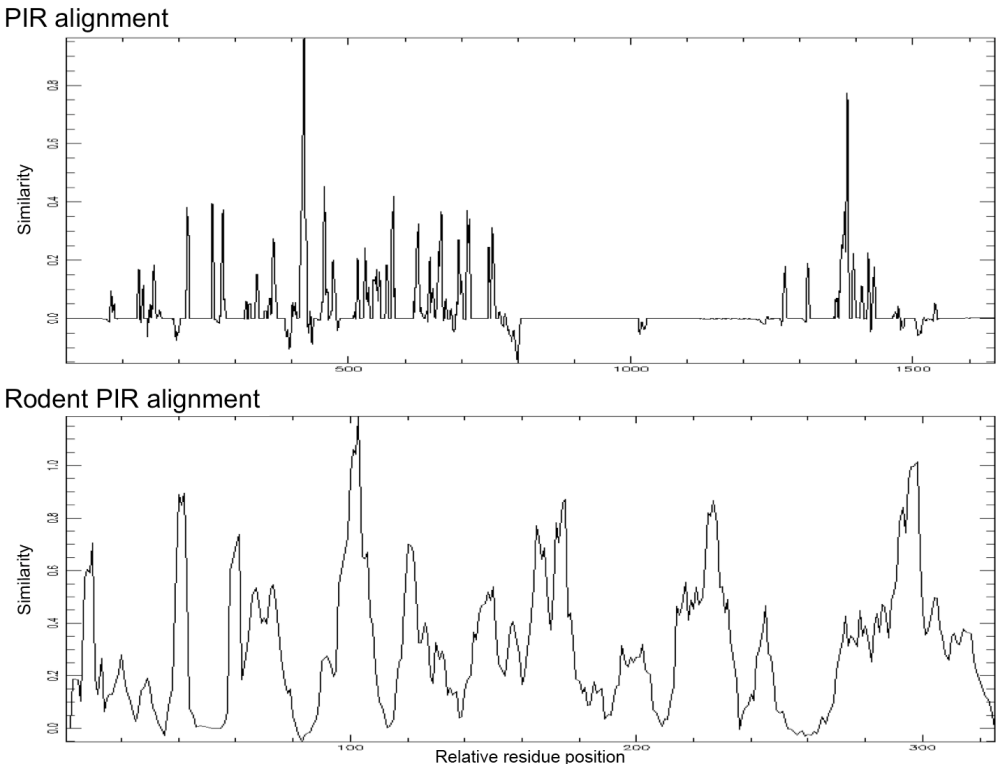
Appendix 2.1 Common *Plasmodium* splice sites

Splice boundary	DNA sequence	Amino acid sequence
Exon 1 / Intron 1	TG / GT	EL or KL or LM or NV / -
Intron 1 / Exon 2	AG / TG or AG / TA	- / CG or CK or CD or CE
Exon 2 / Intron 2	TG / GT	YK / -
Intron 2 / Exon 3	AG / TA	- / YSLF

The sequences above are commonly found at the boundaries of *Plasmodium* exons and introns (personal communication, Dr U. Böhme). During *cir* gene annotation (chapter 2.2.1), where these sequences flanked regions without stop codons, this indicated the site of a potential ORF.

Appendix 2.2 Comparison of PIR alignments

136 PIR and 500 rodent PIR amino acid sequences were aligned using Muscle (Edgar, 2004). Sequence similarity for each alignment was plotted using Plotcon (http://emboss.bioinformatics.nl/cgi-bin/emboss/plotcon). Gaps present in the alignment are indicated by low sequence similarity.



The following appendices for chapter 2 were not suitable for print due to their large size, however they are enclosed within the CD accompanying this thesis. Appendices 2.2, 2.5, 2.7, 2.11 and 2.13 are FASTA files, which may be viewed using software such as Bioedit (Hall, 1999). The remaining appendices for chapter 2 are TIFF or PDF files, and may be viewed using software such as Adobe Photoshop.

Appendix 2.3 Alignment of 107 CIR amino acid sequences.

Appendix 2.4 Maximum likelihood tree of 107 CIRs.

Appendix 2.5 NeighborNet network of 107 CIRs.

Appendix 2.6 Alignment of 117 CIR amino acid sequences.

Appendix 2.7 Maximum likelihood tree of 117 CIRs.

Appendix 2.8 NeighborNet network of 117 CIRs.

Appendix 2.9 Alignment of 183 CIR amino acid sequences.

Appendix 2.10 Maximum likelihood tree of 183 CIRs.

Appendix 2.11 NeighborNet network of 183 CIRs.

Appendix 2.12 Alignment of 136 PIR sequences.

Appendix 2.13 NeighborNet network of 136 PIRs.

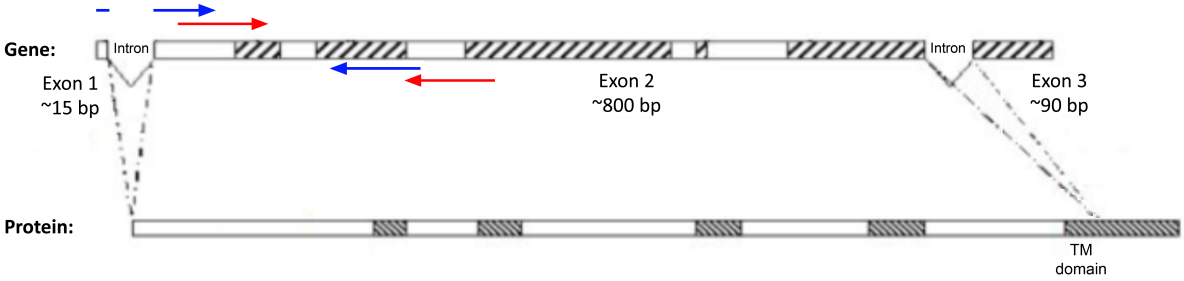
Appendix 2.14 Alignment of 500 rodent PIR sequences.

Appendix 2.15 NeighborNet network of 500 rodent PIRs.

Appendix 2.16 Significant CIR CSS sites, where Z > 0.5.

Appendix 2.17 Significant CIR RSS sites, where U > 4.

Appendix 3.1 *cir* primer design



Appendix 3.1.1 Location of *cir* primers within a typical *cir* gene

This figure illustrates a typical *cir* gene and its encoded protein, with regions of sequence conservation highlighted by hashed lines. In order to amplify a small number of *cir* genes from different sub-families with different primer pairs, primer design was focussed within the variable sequence towards the 5' end of *cir* genes.

Two strategies were initially employed for *cir* primer design: The first, indicated above by blue arrows, used a forward primer that spanned an exon: exon boundary. These primers are listed in Appendix 3.1.2. The second, indicated above by red arrows, used both primers situated within the *cir* second exon. These primers are listed in Appendix 3.1.3

Appendix 3.1.2 Strategy 1) Design of exon: exon boundary spanning primers

ID	Fwd. sequence	Rev. sequence (RC)	BLAST matches	Predicted dimers	Product size	Tm
A	ATGGCTGAAAATATGTGTAGAAG	ACATTGTCGATTTGGCCATC	PCHAS_073190 PCHAS_100030	8	296	58.7 °C 64.7 °C
B	ATGTCTAAGGGAGTGTGTGAAG	GCTCGAACATCACTAGCAACC	PCHAS_000410 PCHAS_000320	8	202	59.7 °C 63.7 °C
B	"	TGTCCATTTCTCCCATTGTTC	PCHAS_000410 PCHAS_000320	6	140	59.7 °C 63.6 °C
C	ATGTCTAAGAAAGTGTGTGAAGC	ATTTTTCGGGATAGGGCAAT	PCHAS_120020 PCHAS_130020	5	126	59.0 °C 63.1 °C
D	CTGCAGTGTGTAACGCAATTAAAG	ATCACCATCGTAAACATCAACAAGAT	PCHAS_070130	8	217	64.4 °C 65.2 °C
D	ATGGCTTATGGAGTGTGTAACG	GCGTATTCGGCAAGTTTATCA	PCHAS_120070 PCHAS_001130	4	251	62.6 °C 63.7 °C
D	"	TCAATGCATGGTTAGGGGTAA	PCHAS_120070 PCHAS_001130	4	100	62.6 °C 63.9 °C
E	TAAGATTGAGGATGTGTATAATGA	TCTTTTTCTTGTCTATCGTCGTAATG	PCHAS_000100	7	137	56.8 °C 63.2 °C
E	CAAGTTATAAGATTGAGGATGTGTATAAAG	GCGTATTCGGCAAGTTTATCAT	PCHAS_120060 PCHAS_000420 PCHAS_000310	7	252	61.1 °C 63.7 °C
U	GTCTAAGGAATTGTGTCAAGGAA	TCTATGGATTTACCATCTTCTCC	PCHAS_040110	5	341	60.9 °C 59.9 °C
U	ATGTCTGAGGAATTGTGTGGAG	TTCCTTTTCCATCCTTTCCA	PCHAS_120030 PCHAS_030030	7	85	62.9 °C 63.2 °C

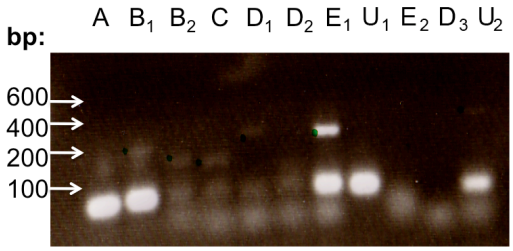
Primer design was described in chapter 3. The *cir* sub-family targeted by the primers is indicated in the ID column. The forward (Fwd) and reverse (Rev) sequences are given for each primer pair, where the reverse primer sequence shown is the reverse complement (RC) of that present in the corresponding *cir* gene. Tm refers to the melting temperature of each primer. The product size is shown as length in nucleotides.

Appendix 3.1.3 Strategy 2) Design of primers within the *cir* second exon

Strategy 2) Design of sub-group specific non-intron spanning primers						
ID	**Fwd. sequence**	**Rev. sequence (RC)**	**BLAST matches**	**Predicted dimers**	**Product size**	**Tm**
A	GGTGAGGGGTATGAAACGTC	TGCAATCTGGTGGGTCACTA	PCHAS_000270 PCHAS_000090 PCHAS_000030	2	180	63.2 °C 64.2 °C
A	CACCAGATTGCACCAAACAT	GGGGGATATTGTCACAACCA	PCHAS_000270 PCHAS_000090 PCHAS_000030 PCHAS_070030 PCHAS_011520	3	188	63.5 °C 64.4 °C
B	CATTCCTTAATAATAATGATGTTG	GTCGTTGGAAGGAGTGGAAA	PCHAS_104250 PCHAS_000410 PCHAS_000320	3	138	56.0 °C 64.0 °C
B	TGGATGCAGCAATATTCCAA	GCAATCCCCAAGAAAAGTGA	PCHAS_000130	2	236	63.9 °C 63.9 °C
B	"	TGAACCCGTTTACCAAATCC	PCHAS_000130	7	197	63.9 °C 63.4 °C
B	TGCGAGTTTCCATCTCTTCCA	TGCCCCTCTATCTTTGTAGGG	PCHAS_000400 PCHAS_000330 PCHAS_040020	4	226	67.0 °C 64.0 °C
B	CTGTTCGCAAAATGCTGAAG	CATGACTTGGCGTAGAACCTT	PCHAS_000400 PCHAS_000080 PCHAS_001070	4	227	63.5 °C 63.0 °C
C	GAATGCACCGAACTTCCAAC	GCAAATCCCAAGAAAATTGG	PCHAS_114750 PCHAS_000140	5	142	64.4 °C 63.0 °C
C	CTTCCACCACTTCCAACGAT	GAGCCCGTTTATCAAATCCA	PCHAS_130020	1	237	63.9 °C 63.6 °C
C	CGAGATGCTCAAGATTTTGCT	CGATGAACTTGGTGGTGCTT	PCHAS_030030 PCHAS_000350 PCHAS_110020	4	249	63.3 °C 65.0 °C
C	CAAGGATGAGGAAGGCACAT	GCTTCACCAGACATTTTCTCAA	PCHAS_030030	6	198	64.0 °C 63.1 °C
D	CGCAATTAAAGCGATCGATAA	TCCCTTTTTGTGGGTTTTGA	PCHAS_070130	7	280	63.6 °C 64.1 °C
D	TTACCCCTAACCATGCATTGA	GCGTATTCGGCAAGTTTATCA	PCHAS_120070 PCHAS_000300 PCHAS_001130 PCHAS_100040	5	171	63.9 °C 63.7 °C
D	ACCCCTAACCATGCATTGAA	GCGTATTCGGCAAGTTTAGC	PCHAS_001130 PCHAS_120070 PCHAS_100040	3	169	64.0 °C 63.5 °C
D	GGCTTCTGCAGTGTGTGGT	GCGTATTCGGCAAGTTTAGC	PCHAS_100040	2	248	64.2 °C 63.5 °C
E	TTGATGAGGATGAAAATGATGG	CAAATGAATAACACCAGAACTAGTCA	PCHAS_000100	6	140	63.5 °C 62.0 °C
E	CATCCCTCCCAGTTTATCCA	TGGAATAATCCCTCTCTTCGAG	PCHAS_000100 PCHAS_000420	6	218	63.6 °C 63.2 °C
E	TGGTGGTGATGGTCTGACTT	TTTTCGCAATCCGGATGT	PCHAS_000340 PCHAS_104230 PCHAS_030040 PCHAS_040040	0	152	63.2 °C 64.1 °C
E	ACATCCGGATTGCGAAAA	TGGATAAACTGGGAGGGATG	PCHAS_000340 PCHAS_104230 PCHAS_030040 PCHAS_040040	3	177	64.1 °C 63.6 °C
U	TGGCCTAAAAATTGGCTCTG	TCTATGGATTTACCATCTTCTCC	PCHAS_040110	-2.16	213	63.9 °C 59.9 °C
U	AACGTCGTTTTGGACTCAGC	CATCTTCGCAGTTCCCCTAA	PCHAS_040050 PCHAS_001050	3	233	64.2 °C 64.0 °C
U	GTGCTGCTTCGTATCGTTCA	CTGCCTTCGGTGCGTATATT	PCHAS_070100	1	249	64.0 °C 63.8 °C
U	"	TTCCTTCTGGTGGCTTCAAT	PCHAS_070100	0	149	64.0 °C 63.5 °C
U	CTGACTTGTATCGCGCATGT	ACACTGGCTTTATGCCGTTC	PCHAS_042030 PCHAS_040110 PCHAS_030210 PCHAS_060060	3	199	63.9 °C 63.9 °C
U	GAACGGCATAAAGCCAGTGT	TCGCCATACGGGTGTTAAAT	PCHAS_040230 PCHAS_146770 PCHAS_030210 PCHAS_060060 PCHAS_060050 PCHAS_040110	4	207	63.9 °C 63.9 °C

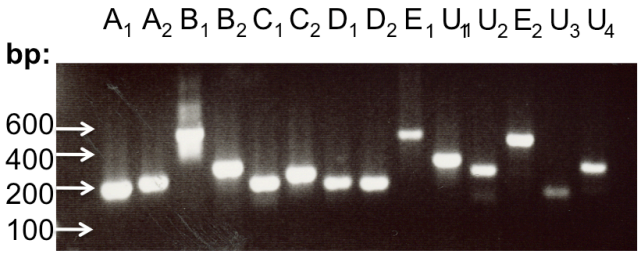
Primer design was described in chapter 3. The *cir* sub-family targeted by the primers is indicated in the ID column. The forward (Fwd) and reverse (Rev) sequences are given for each primer pair, where the reverse primer sequence shown is the reverse complement (RC) of that present in the corresponding *cir* gene. Tm refers to the melting temperature of each primer. The product size is shown as length in nucleotides.

Appendix 3.2 *cir* primer efficacy tested using PCR



Appendix 3.2.1 Strategy 1) exon: exon boundary spanning primers

This figure shows the amplification products of primers listed in Appendix 3.1.2, after endpoint PCR. Only one of the primers from strategy 1 (D2) showed marked amplification of any product. Primers designed using strategy 1 could thus either poorly amplify their target genes or the target *cir* genes may not have been expressed in the cDNA template. Since it would be difficult and lengthy to investigate which of these was the case, and in addition, extensive primer dimer formation (bands less than 100 bp in size) was observed for most primers, the use of all strategy 1 primers was discontinued at this point.



Appendix 3.2.2 Strategy 2) primers within *cir* exons

This figure shows the amplification products of primers listed in Appendix 3.1.3, after endpoint PCR. The majority of primers from strategy 2 amplified products of the expected size, between 150 and 250 bp. Primers that successfully amplified products are listed in Appendix 3.2.3

Appendix 3.2.3 Primers that amplified products by endpoint PCR.

ID	Fwd. sequence	Rev. sequence (RC)	BLAST matches	Predicted dimers	Product size	Tm
A1	GGTGAGGGGTATGAAACGTC	TGCAATCTGGTGGGTCACTA	PCHAS_000270 PCHAS_000090 PCHAS_000030	2	180	63.2 °C 64.2 °C
A2	CACCAGATTGCACCAAACAT	GGGGGATATTGTCACAACCA	PCHAS_000270 PCHAS_000090 PCHAS_000030 PCHAS_070030	3	188	63.5 °C 64.4 °C
B1	TGCGAGTTTCCATCTCTTCCA	TGCCCCTCTATCTTTGTAGGG	PCHAS_000400 PCHAS_000330 PCHAS_040020	4	226	67.0 °C 64.0 °C
B2	CTGTTCGCAAAATGCTGAAG	CATGACTTGGCGTAGAACCTT	PCHAS_000400 PCHAS_000080 PCHAS_001070	4	227	63.5 °C 63.0 °C
C1	GAATGCACCGAACTTCCAAC	GCAAATCCCAAGAAAATTGG	PCHAS_114750 PCHAS_000140	5	142	64.4 °C 63.0 °C
C2	CGAGATGCTCAAGATTTTGCT	CGATGAACTTGGTGGTGCTT	PCHAS_030030 PCHAS_000350 PCHAS_110020	4	249	63.3 °C 65.0 °C
D1	TTACCCCTAACCATGCATTGA	GCGTATTCGGCAAGTTTATCA	PCHAS_120070 PCHAS_000300 PCHAS_001130 PCHAS_100040	5	171	63.9 °C 63.7 °C
D2	ACCCCTAACCATGCATTGAA	GCGTATTCGGCAAGTTTAGC	PCHAS_001130 PCHAS_120070 PCHAS_100040	3	169	64.0 °C 63.5 °C
E1	TTGATGAGGATGAAAATGATGG	CAAATGAATAACACCAGAACTAGTCA	PCHAS_000100	6	140	63.5 °C 62.0 °C
E2	CATCCCTCCCAGTTTATCCA	TGGAATAATCCCTCTCTTCGAG	PCHAS_000100 PCHAS_000420	6	218	63.6 °C 63.2 °C
U1	AACGTCGTTTTGGACTCAGC	CATCTTCGCAGTTCCCCTAA	PCHAS_040050 PCHAS_001050	3	233	64.2 °C 64.0 °C
U2	TGGCCTAAAAATTGGCTCTG	TCTATGGATTTACCATCTTCTCC	PCHAS_040110	5	213	63.9 °C 59.9 °C
U3	CTGACTTGTATCGCGCATGT	ACACTGGCTTTATGCCGTTC	PCHAS_042030 PCHAS_040110 PCHAS_030210 PCHAS_060060	3	199	63.9 °C 63.9 °C
U4	GAACGGCATAAAGCCAGTGT	TCGCCATACGGGTGTTAAAT	PCHAS_040230 PCHAS_146770 PCHAS_030210 PCHAS_060060 PCHAS_060050 PCHAS_040110	4	207	63.9 °C 63.9 °C

Appendix 3.3 RNA integrity

Three experiments were carried out in chapter 3 to measure *cir* transcript levels during *P. chabaudi* infection by RT-qPCR. All mice were labelled with their cage number:

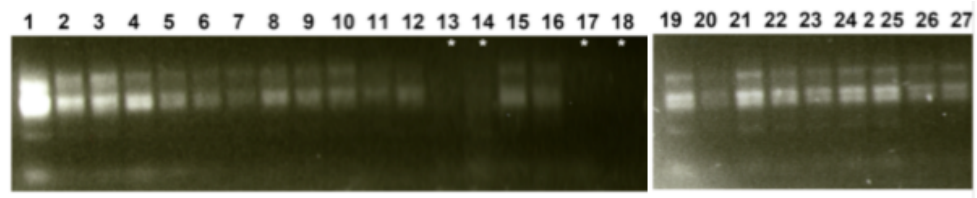
- Figure 22a used *P. chabaudi* parasites expanded in a RAG2$^{-/-}$ mouse (JL1014) to infect three BALB mice (JL1017-1, 2 and 3). As the whole blood volume of these mice was taken, RNA was extracted from two or three aliquots per mouse.
- Figure 22b used *P. chabaudi* parasites from the BALB/c mouse JL1017-1 to infect three RAG2$^{-/-}$ mice (JL1018-1, 2 and 3). As the whole blood volume of these mice was taken, RNA was extracted from two or three aliquots per mouse.
- Figure 23 re-capitulated the experiment described in Figure 22b, using *P. chabaudi* parasites expanded in a BALB/c mouse to infect three RAG2$^{-/-}$ mice (JL1043 M1, M2 and M3). Samples were taken from these mice at 8.30 am, 10.30 am, 12.30 pm and 10.30 pm.

Appendix 3.3.1 Spectroscopy measurements of extracted RNA.

Sample	Tube	Concentration (ng/ul)	500ng	A260/A280
JL1014 RAG	1	213.63	2.34	1.92
	* 2	413.23	1.21	2.01
	* 3	232.95	2.15	1.94
JL1017-1 BALB	1	341.09	1.47	1.96
	* 2	413.55	1.21	2.02
JL1017-2 BALB	* 1	592.93	0.84	1.87
	2	499.09	1.00	1.94
JL1017-3 BALB	* 1	510.49	0.98	1.75
	* 2	346.95	1.44	1.99
	3	334.6	1.49	2.00
JL1018-1 RAG	* 1	441.96	1.13	2.07
	2	457.95	1.09	2.10
JL1018-2 RAG	* 1	420.18	1.19	2.10
	2	415.23	1.20	2.08
JL1018-3 RAG	1	591.55	0.85	2.07
	* 2	600.86	0.83	2.09
JL1043 RAG 8-30am	M1	232.02	2.15	1.87
	M2	260.55	1.92	1.91
	M3	334.55	1.49	1.98
JL1043 RAG 10-30am	M1	235.47	2.12	1.86
	M2	326.87	1.53	1.98
	M3	366.42	1.36	1.98
JL1043 RAG 12-30pm	M1	388.31	1.29	2.01
	M2	224.09	2.23	1.88
	M3	564.39	0.89	1.71
JL1043 RAG 10-30pm	M1	665.3	0.75	1.71
	M2	337.65	1.48	1.94
	M3	551.03	0.91	1.72

Samples indicated with a * were pooled to produce a positive control for RT-qPCR (hereafter referred to as the cDNA pool). All other samples used to measure *cir* transcription in individual mice.

Appendix 3.3.2 Gel electrophoresis of 500ng RNA.



Where samples are numbered as follows:

1) JL1014 M1-1	10) JL1018 M1-1	19) JL1043 M1 10.30
2) JL1014 M1-2	11) JL1018 M1-2	20) JL1043 M2 10.30
3) JL1017 M1-1	12) JL1018 M2-1	21) JL1043 M3 10.30
4) JL1017 M1-2	13) JL1018 M2-2	22) JL1043 M1 12.30
5) JL1017 M2-1	14) JL1018 M3-1	23) JL1043 M2 12.30
6) JL1017 M2-2	15) JL1018 M3-2	24) JL1043 M3 12.30
7) JL1017 M3-1	16) JL1043 M1 8.30	25) JL1043 M1 22.30
8) JL1017 M3-2	17) JL1043 M2 8.30	26) JL1043 M2 22.30
9) JL1017 M3-3	18) JL1043 M3 8.30	27) JL1043 M3 22.30

Samples marked with * were precipitated and the concentrations re-measured prior to use for cDNA synthesis (data not shown).

Appendix 3.4 RT-qPCR analysis

Appendix 3.4.1 Process of analysis of raw Cq values from RT-qPCR data.

Cq values obtained from RT-qPCR experiments were analyzed as outlined in chapter 3:

1) Calculation of the mean Cq for each sample from three technical replicates.
2) Calculation of the standard deviation for each mean Cq.
3) Where one replicate caused the SD to increase above the 80% confidence level, this replicate was excluded. The mean and SD of the Cq values were re-calculated from the two technical replicates.
4) The following equation was used to calculate the ratio of change in *cir* transcript levels between the 'control' and 'sample' mice (Pfaffl, 2001):
   $\text{Ratio} = (E_{cir})^{\Delta Cq\ cir\ (\text{control - sample})} / (E_{beta\ tubulin})^{\Delta Cq\ beta\ tubulin\ (\text{control - sample})}$
   Where E represents primer efficiency.
5) The standard deviation of differences in Cq between samples were calculated using the equation: $SD\Delta Cq = \sqrt{[(SD\ beta\ tubulin)^2 + (SD\ beta\ globin)^2]}$

The mean ratio, SD and number of technical replicates (N) used for figures 22 and 23 are shown below, in appendices 3.5.2-4. The confidence interval (CI) is indicated by the following colour code. For example the maximum allowable SD that can distinguish a two-fold change in transcription levels with 99% confidence is 0.3.

SD:	< 0.3	< 0.4	< 0.525	< 0.650	< 0.775	> 0.775
CI:	99%	95%	90%	85%	80%	exclude

Appendix 3.4.2 RT-qPCR data in Figure 21a (compared to donor mouse)

BALB/c mouse	U1			U4			C1			E1		
	Mean ratio	SD	N	Mean ratio	SD	N	Mean ratio	SD	N	Mean ratio	SD	N
1	0.33	0.04	2.00	0.24	0.28	2.00	1.31	0.21	3.00	0.69	0.15	3.00
2	0.61	0.25	2.00	0.51	0.27	3.00	1.52	0.14	3.00	0.68	0.34	2.00
3	0.61	0.10	3.00	0.08	0.32	2.00	1.12	0.28	2.00	0.29	0.41	3.00

Appendix 3.4.3 RT-qPCR data in Figure 21b (compared to donor mouse)

Mouse	U1			U4			C1			E1		
	Mean ratio	SD	N	Mean ratio	SD	N	Mean ratio	SD	N	Mean ratio	SD	N
1	3.49	0.28	2.00	2.58	0.07	2.00	2.37	0.20	2.00	1.05	0.18	2.00
2	1.89	0.36	3.00	2.07	0.12	2.00	0.76	0.18	3.00	1.81	0.25	3.00
3	3.37	0.27	3.00	2.29	0.19	3.00	1.20	0.13	3.00	1.55	0.31	3.00

Appendix 3.4.4 RT-qPCR data in Figure 22 (compared to 8.30h time-point)

Time-point	U1			U4			C1			E1		
	Mean ratio	SD	N	Mean ratio	SD	N	Mean ratio	SD	N	Mean ratio	SD	N
10.30 M1	2.12	0.04	3.00	6.15	0.30	2.00	2.15	0.17	3.00	1.50	0.40	3.00
10.30 M2	2.30	0.19	2.00	2.63	0.43	2.00	1.36	0.11	2.00	3.57	0.41	2.00
10.30 M3	6.79	0.55	2.00	3.91	0.17	2.00	1.96	0.27	3.00	3.00	0.27	3.00
12.30 M1	5.42	0.17	3.00	5.59	0.46	2.00	3.22	0.08	2.00	3.38	0.25	3.00
12.30 M2	19.13	0.09	3.00	0.21	0.32	2.00	9.27	0.23	3.00	6.32	0.08	3.00
12.30 M3	1.54	0.37	2.00	2.68	0.04	2.00	0.55	0.26	2.00	1.36	0.37	3.00
22.30 M1	2.51	0.16	3.00	12.27	0.25	2.00	2.68	0.12	2.00	13.62	0.31	2.00
22.30 M2	5.15	0.22	3.00	6.81	0.30	2.00	4.88	0.08	2.00	23.17	0.24	2.00
22.30 M3	3.35	0.51	3.00	6.84	0.60	2.00	7.01	0.34	2.00	10.04	0.15	3.00

Appendix 3.5 RNA sequencing data

Appendix 3.5.1 Sample 1

For each *cir* gene, the total length is given, followed by length of the gene found to be unique to that *cir*, or present in other *cir*s. The number of RNA sequence reads that were mapped to each *cir* gene is given, along with the mean, geometric mean, median and the minimum and maximum errors of the mean. In addition, the product of each gene is given, as transcripts belonging to three highly divergent '*cir*-like' genes were also detected.

Name	Gene length	Non-unique sequence	Unique sequence	Number of reads mapping	Max	Min	Geometric Mean	Mean	Median	Product
PCAS_010020	893	380	513	6	2	0	1.05	0.4	1	CIR protein
PCAS_010030	902	671	231	2	2	0	1	0.39	1	CIR protein
PCAS_010120	1055	509	546	1100	155	8	67.25	74.39	75	CIR protein
PCAS_011450	890	0	890	50	11	0	1.7	2.06	2	CIR protein
PCAS_011510	959	57	902	5	2	0	1.02	0.22	1	CIR protein
PCAS_011520	917	0	917	6	2	0	1.03	0.23	1	CIR protein
PCAS_011530	902	38	864	6	1	0	1	0.26	1	CIR protein
PCAS_030020	881	0	881	14	2	0	1.07	0.6	1	CIR protein
PCAS_030030	845	272	573	4	1	0	1	0.24	1	CIR protein
PCAS_030040	917	380	537	19	4	0	1.5	1.33	2	CIR protein
PCAS_030060	953	514	439	0	0	0	1	0	1	CIR protein
PCAS_030080	959	284	675	2	2	0	1.01	0.09	1	CIR protein
PCAS_030090	950	24	926	30	6	0	1.39	1.19	2	CIR protein
PCAS_030110	941	796	145	5	4	0	1.44	1.31	2	CIR protein
PCAS_030120	1079	922	157	0	0	0	1	0	1	CIR protein
PCAS_030150	1742	341	1401	2	1	0	1	0.05	1	CIR protein
PCAS_030180	2030	0	2030	2	1	0	1	0.04	1	CIR protein
PCAS_030220	959	0	959	37	6	0	1.54	1.42	2	CIR protein
PCAS_030210	959	0	959	37	6	0	1.54	1.42	2	CIR protein
PCAS_030270	1301	0	1301	116	9	0	2.77	3.31	4	CIR protein
PCAS_040010	941	315	626	20	6	0	1.4	1.18	2	CIR protein
PCAS_040020	1718	1480	238	3	2	0	1.03	0.44	1	CIR protein
PCAS_040030	947	35	912	15	7	0	1.1	0.6	1	CIR protein
PCAS_040040	914	573	341	14	6	0	1.6	1.56	2	CIR protein
PCAS_040050	980	257	723	6	2	0	1.04	0.31	1	CIR protein
PCAS_040110	941	214	727	22	5	0	1.32	1.09	2	CIR protein
PCAS_042030	950	86	864	3	2	0	1.01	0.13	1	CIR protein
PCAS_050020	932	0	932	60	17	0	1.73	2.4	2	CIR protein
PCAS_050060	857	0	857	16	4	0	1.15	0.68	1	CIR protein
PCAS_050070	917	16	901	12	3	0	1.1	0.48	1	CIR protein
PCAS_060040	950	0	950	14	3	0	1.08	0.54	1	CIR protein
PCAS_070020	974	487	487	3	3	0	1.06	0.25	1	CIR protein
PCAS_070030	932	478	454	6	2	0	1.07	0.45	1	CIR protein
PCAS_070050	893	465	428	11	6	0	1.22	0.98	2	CIR protein
PCAS_070100	1073	30	1043	26	5	0	1.24	0.92	2	CIR protein
PCAS_070130	893	50	843	166	20	0	6.23	7.29	8	CIR protein
PCAS_070160	902	41	861	4	2	0	1.02	0.16	1	CIR protein
PCAS_073160	896	0	896	15	5	0	1.14	0.62	1	CIR protein
PCAS_073180	932	0	932	15	5	0	1.11	0.6	1	CIR protein
PCAS_073190	905	0	905	14	4	0	1.11	0.59	1	CIR protein
PCAS_073200	929	0	929	51	8	0	1.85	2.02	3	CIR protein
PCAS_093820	731	717	14	0	0	0	1	0	1	CIR protein
PCAS_104000	1124	0	1124	92	11	0	2.39	3.03	3	CIR protein
PCAS_104110	917	0	917	15	3	0	1.13	0.61	1	CIR protein
PCAS_104100	917	0	917	15	3	0	1.13	0.61	1	CIR protein
PCAS_104120	920	0	920	38	12	0	1.49	1.53	2	CIR protein
PCAS_104130	926	0	926	29	12	0	1.39	1.18	1	CIR protein
PCAS_110020	845	0	845	108	18	0	3.73	4.72	5	CIR protein
PCAS_120020	929	0	929	17	6	0	1.17	0.66	1	CIR protein
PCAS_120030	953	0	953	25	4	0	1.28	0.97	2	CIR protein
PCAS_120040	986	0	986	8	2	0	1.03	0.31	1	CIR protein
PCAS_120050	914	0	914	8	2	0	1.04	0.34	1	CIR protein
PCAS_120060	911	0	911	18	4	0	1.26	0.72	1	CIR protein
PCAS_120070	881	0	881	11	3	0	1.05	0.44	1	CIR protein
PCAS_124630	1106	513	593	11	4	0	1.12	0.69	1	CIR protein
PCAS_124690	1100	0	1100	11	3	0	1.06	0.36	1	CIR protein
PCAS_130020	959	0	959	12	6	0	1.11	0.48	1	CIR protein
PCAS_130030	950	0	950	9	2	0	1.04	0.34	1	CIR protein
PCAS_130050	944	0	944	4	1	0	1	0.15	1	CIR protein
PCAS_130070	1985	0	1985	9	2	0	1.02	0.16	1	CIR protein
PCAS_130080	2318	0	2318	11	3	0	1.01	0.18	1	CIR protein

PCAS_137120	923	815	108	1	1	0	1	0.28	1	CIR protein
PCAS_137130	911	447	464	5	2	0	1.03	0.37	1	CIR protein
PCAS_137150	1058	0	1058	19	4	0	1.14	0.66	1	CIR protein
PCAS_137160	878	0	878	106	11	0	3.61	4.48	5	CIR protein
PCAS_140020	968	0	968	20	6	0	1.18	0.76	1	CIR protein
PCAS_140030	917	0	917	12	4	0	1.15	0.49	1	CIR protein
PCAS_140040	926	0	926	18	4	0	1.14	0.7	1	CIR protein
PCAS_140070	929	0	929	28	4	0	1.3	1.12	2	CIR protein
PCAS_140090	860	0	860	6	4	0	1.01	0.28	1	CIR protein
PCAS_140130	866	0	866	43	8	0	1.68	1.84	2	CIR protein
PCAS_146850	1082	0	1082	13	3	0	1.08	0.45	1	CIR protein
PCAS_146870	935	0	935	3	2	0	1	0.12	1	CIR protein
PCAS_052430	716	0	716	978	211	4	35.81	50.46	45	CIR-like protein
PCAS_090030	1199	0	1199	36	8	0	1.32	1.1	2	CIR-like protein
PCAS_146840	587	0	587	22	9	0	1.45	1.38	1	CIR-like protein

Appendix 3.5.2 Sample 2

Here, only the number of reads mapping to each *cir* gene is given, as full statistics for this sample had not been calculated at the time of writing.

Name	Number of reads mapping	Name	Number of reads mapping	Name	Number of reads mapping
PCHAS_040030	124.4815679	PCHAS_011510	0	PCHAS_114700	13.20259054
PCHAS_040060	299.8874136	PCHAS_011520	49.03819342	PCHAS_114720	70.72816359
PCHAS_000320	0	PCHAS_011530	1.886084362	PCHAS_114750	0
PCHAS_070130	794.9845588	PCHAS_030020	1.886084362	PCHAS_120020	0
PCHAS_070160	0	PCHAS_030040	0	PCHAS_120040	0
PCHAS_041970	5.658253087	PCHAS_030070	1.886084362	PCHAS_120060	0.943042181
PCHAS_090030	120.7093992	PCHAS_030090	4.715210906	PCHAS_120070	1.886084362
PCHAS_060030	33.94951852	PCHAS_030110	3.772168725	PCHAS_130030	0
PCHAS_073180	76.38641668	PCHAS_030120	7.54433745	PCHAS_130060	1.886084362
PCHAS_000580	0	PCHAS_030140	0	PCHAS_130070	0
PCHAS_001120	0	PCHAS_030160	7.54433745	PCHAS_130080	24.51909671
PCHAS_000020	0	PCHAS_030190	0	PCHAS_130090	59.41165742
PCHAS_000040	9.430421812	PCHAS_030200	16.97475926	PCHAS_130100	1.886084362
PCHAS_000060	7.54433745	PCHAS_030210	9.430421812	PCHAS_130110	8.487379631
PCHAS_000070	9.430421812	PCHAS_030270	193.3236472	PCHAS_130120	15.0886749
PCHAS_000090	0	PCHAS_040080	3.772168725	PCHAS_130170	7.54433745
PCHAS_000100	22.63301235	PCHAS_040110	73.55729014	PCHAS_130220	1.886084362
PCHAS_000110	9.430421812	PCHAS_041950	0	PCHAS_130280	1.886084362
PCHAS_000120	12.25954836	PCHAS_041980	0	PCHAS_137030	58.46861524
PCHAS_000130	9.430421812	PCHAS_041990	0	PCHAS_137090	1.886084362
PCHAS_000140	0	PCHAS_042020	0	PCHAS_137110	0
PCHAS_000150	8.487379631	PCHAS_042030	5.658253087	PCHAS_140030	3.772168725
PCHAS_000170	1.886084362	PCHAS_042070	23.57605453	PCHAS_140040	1.886084362
PCHAS_000220	0	PCHAS_050020	5.658253087	PCHAS_140070	5.658253087
PCHAS_000270	0	PCHAS_050040	33.94951852	PCHAS_140090	3.772168725
PCHAS_000310	1.886084362	PCHAS_050060	1.886084362	PCHAS_140140	10.37346399
PCHAS_000340	0	PCHAS_050070	3.772168725	PCHAS_146770	7.54433745
PCHAS_000350	0.943042181	PCHAS_060060	0	PCHAS_146840	96.19030249
PCHAS_000360	0	PCHAS_060070	13.20259054	PCHAS_146860	0
PCHAS_000400	0.943042181	PCHAS_060100	20.74692799	PCHAS_146870	0
PCHAS_000410	0	PCHAS_060110	30.1773498	PCHAS_000180	0
PCHAS_000420	1.886084362	PCHAS_060120	102.7915978	PCHAS_041960	0
PCHAS_000470	9.430421812	PCHAS_060130	15.0886749	PCHAS_042010	0
PCHAS_000490	12.25954836	PCHAS_060160	1.886084362	PCHAS_060080	0
PCHAS_000500	11.31650617	PCHAS_070020	11.31650617	PCHAS_140130	0
PCHAS_000560	1.886084362	PCHAS_070030	1.886084362	PCHAS_001010	0
PCHAS_000570	0	PCHAS_070040	52.81036215	PCHAS_030150	0
PCHAS_000680	0	PCHAS_070070	11.31650617	PCHAS_040040	18.86084362
PCHAS_000720	8.487379631	PCHAS_070100	10.37346399	PCHAS_070090	14.14563272
PCHAS_000730	3.772168725	PCHAS_070170	1.886084362	PCHAS_073200	25.46213889
PCHAS_000770	2.829126544	PCHAS_073150	5.658253087	PCHAS_104250	0
PCHAS_000950	0	PCHAS_073160	0.943042181	PCHAS_130350	18.86084362
PCHAS_001040	5.658253087	PCHAS_073190	32.06343416	PCHAS_140020	18.86084362
PCHAS_001050	0	PCHAS_083690	370.6155772	PCHAS_030030	15.0886749
PCHAS_001060	0.943042181	PCHAS_083720	18.86084362	PCHAS_030060	0.943042181
PCHAS_001090	126.3676523	PCHAS_083760	0	PCHAS_030170	22.63301235
PCHAS_001100	35.83560289	PCHAS_090010	13.20259054	PCHAS_060140	16.97475926
PCHAS_001110	3.772168725	PCHAS_100040	0	PCHAS_070050	0
PCHAS_001130	0	PCHAS_100060	5.658253087	PCHAS_070060	49.03819342
PCHAS_010020	1.886084362	PCHAS_100110	1.886084362	PCHAS_073130	30.1773498
PCHAS_010040	0	PCHAS_104230	0	PCHAS_093820	17.91780144
PCHAS_010080	0.943042181	PCHAS_104260	13.20259054	PCHAS_104200	37.72168725
PCHAS_011330	26.40518107	PCHAS_104270	5.658253087	PCHAS_120050	1.886084362
PCHAS_011450	79.21554322	PCHAS_110020	115.0511461	PCHAS_130050	1.886084362
PCHAS_011480	7.54433745	PCHAS_110030	4238.974605	PCHAS_146790	0
PCHAS_011490	10.37346399	PCHAS_114600	13.20259054	PCHAS_146850	0.943042181
PCHAS_011500	0.943042181	PCHAS_114640	14.14563272		

Appendix 3.6 *Beta globin* primers and RT-qPCR analysis

Appendix 3.6.1 *Beta globin* primers

ID	Sequence	Tm	BLAST matches	Product size
HBB-B2 F	CTGCATGTGGATCCTGAGAA	64.0 °C	NM_016956.2 (mus musculus hemoglobin, beta adult minor chain (Hbb-b2), mRNA) NM_008220.2 (Mus musculus	230 (cDNA) 858 (gDNA)
HBB-B2 R (RC)	TCATTTCCCCACACTTGACA	64.1 °C		

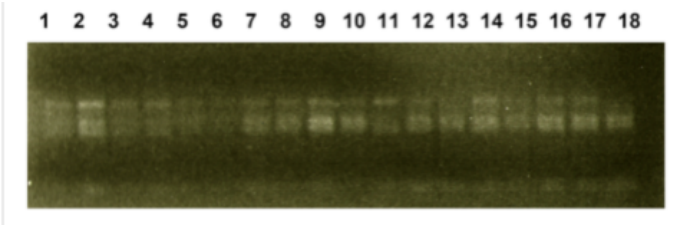
Primer design was described in chapter 3. The ID column indicates the names of the forward (F) and reverse (R) primers which successfully amplified *beta globin* transcripts. The reverse primer sequence is given as the reverse complement (RC) of that present in the *beta globin* gene. Tm refers to the melting temperature of each primer. The product size is shown as length in nucleotides.

Appendix 3.6.2 Analysis of RNA quality by spectroscopy

Sample	Mouse	Concentration (ng/ul)	500ng	A260/A280
10e8 infection	1	364.01	1.37	1.95
	2	591.4	0.85	2.06
	3	841.92	0.59	2.17
10e5 infection	1	258.75	1.93	1.88
	2	228.71	2.19	1.66
	3	339.24	1.47	1.97
MACS	1 before	397.31	1.26	1.72
	1 after	205.2	2.44	2.03
	2 before	861.83	0.58	1.69
	2 after	251.27	1.99	1.70
	3 before	711.78	0.70	1.67
	3 after	295.84	1.69	1.94
Saponin	1 before	435.32	1.15	2.01
	1 after	433.79	1.15	1.97
	2 before	131.23	3.81	1.75
	2 after	379.58	1.32	1.91
	3 before	928.02	0.54	1.82
	3 after	545.3	0.92	1.98

Where mice infected with 10^5 and 10^8 iRBCs were labelled as 10e5 and 10e8, respectively.

Appendix 3.6.3 Analysis of RNA quality by gel electrophoresis



Samples were numbered as follows:

1) $1x10^5$ *P. chabaudi* infection: M1
2) $1x10^5$ *P. chabaudi* infection: M2
3) $1x10^5$ *P. chabaudi* infection: M3
4) $1x10^8$ *P. chabaudi* infection: M1
5) $1x10^8$ *P. chabaudi* infection: M2
6) $1x10^8$ *P. chabaudi* infection: M3
7) Saponin lysis before: M1
8) Saponin lysis before: M2
9) Saponin lysis before: M3
10) Saponin lysis after: M1
11) Saponin lysis after: M2
12) Saponin lysis after: M3
13) MACS purification before: M1
14) MACS purification after: M1
15) MACS purification before: M2
16) MACS purification after: M2
17) MACS purification before: M3
18) MACS purification after: M3

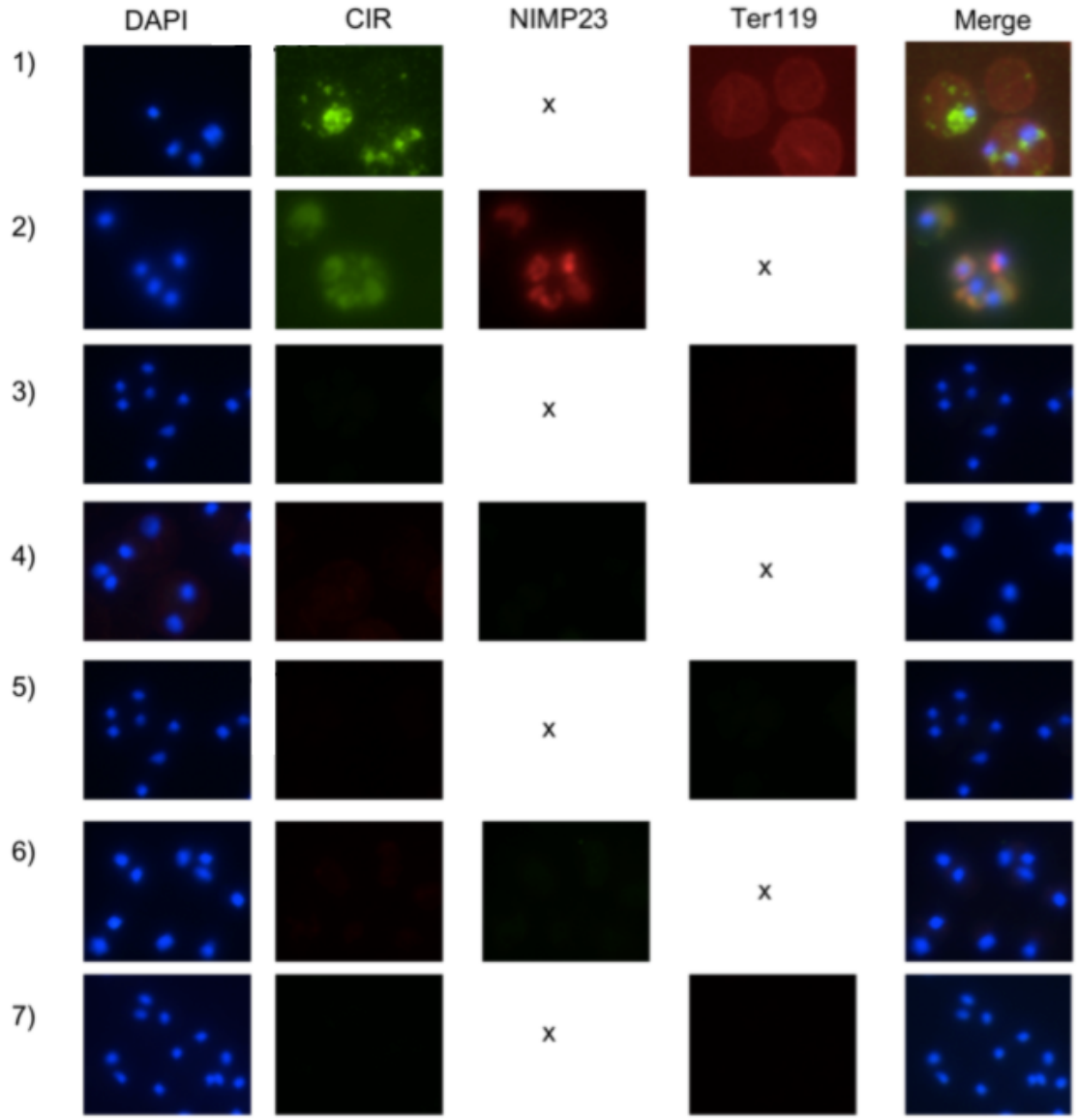
Appendix 3.6.4 Analysis of RT-qPCR data

Cq values obtained from RT-qPCR experiments were analyzed by the following procedure, as outlined in chapter 3 and appendix 3.4.1, with the following exception: The reciprocal mean Cq of *beta globin* was divided by the reciprocal mean Cq of *beta tubulin* for each sample to create a ratio of *beta globin*: *beta tubulin* transcript levels.

Appendix 3.6.5 *beta globin*: *beta tubulin* transcript levels in Figure 24.

Sample	Mouse 1			Mouse 2			Mouse 3		
	Mean ratio	**SD**	**N**	**Mean ratio**	**SD**	**N**	**Mean ratio**	**SD**	**N**
Differential dose experiment. Mice infected with:									
$1x10^5$ iRBCs	1.42	0.13	3	1.42	0.05	3	1.66	0.12	3
$1x10^8$ iRBCs	1.11	0.07	3	1.35	0.17	3	1.28	0.14	3
Saponin lysis experiment:									
Before	1.46	0.18	3	1.24	0.22	3	1.45	0.18	3
After	1.40	0.61	2	1.48	0.01	3	1.41	0.08	3
MACS purification experiment:									
Before	1.36	0.13	3	1.40	0.24	3	1.33	0.19	3
After	1.33	0.11	3	1.43	0.10	2	1.29	0.08	3

Appendix 5.1 Controls for immuno-fluorescence assays

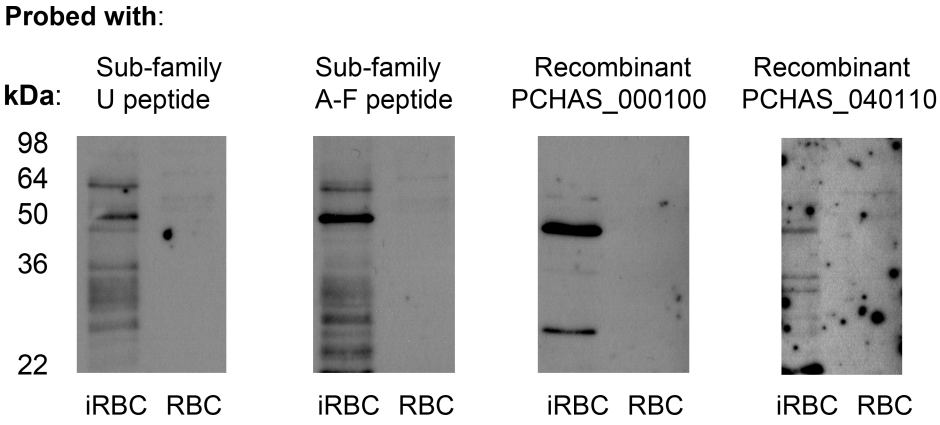


Where:

1) Primary Abs were Anti-CIR & Anti-Rabbit Alexa 488, with the secondary reagents Anti-Glycophorin A (Ter119)-biotin & Streptavidin-Texas Red
2) Primary Abs were Anti-CIR & Anti-Rabbit Alexa 488, with the secondary reagents Anti-MSP1 (NIMP23) & Anti-Mouse Alexa 595
3) No primary antibody, with the secondary reagents Anti-Rabbit Alexa 488 & Streptavidin-Texas Red
4) No primary antibody, with the secondary reagents Anti-Rabbit Alexa 488 & Anti-Mouse Alexa 595
5) Primary Abs were Anti-Glycophorin A (Ter119)-biotin only, with the secondary reagents Anti-Rabbit Alexa 488 & Anti-Mouse Alexa 595
6) Primary Abs were Anti-MSP1 (NIMP23) only with the secondary reagents Anti-Rabbit Alexa 488 & Streptavidin-Texas Red
7) Primary Abs were Anti-CIR only with the secondary reagent Streptavidin-Texas Red

Appendix 6.1 Rationale for the use of CIR peptides and proteins in the immunization and *P. chabaudi* challenge experiment.

Rabbits were immunized with the two CIR sub-family specific peptides, and the recombinant CIR proteins PCHAS_040110 and PCHAS_000100, as described for the conserved CIR peptide in chapter 5. To ensure that Abs generated in response to CIR immunization recognized *P. chabaudi* proteins, rabbit anti-sera were absorbed against uninfected RBC membranes and used to probe western blots (as described in chapter 5).



The figure above shows that all rabbits immunized with CIR proteins or peptides made Abs that recognized *P. chabaudi* proteins (present in the iRBC lanes), which did not recognize proteins from uninfected RBCs. Immunization of mice was thus expected to produce Abs with similar reactivities.

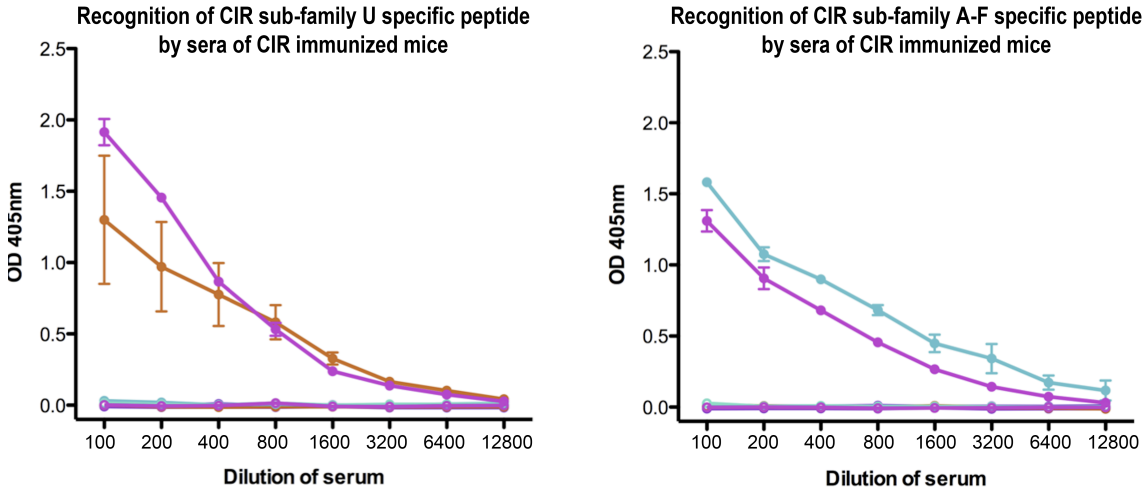
Appendix 6.2 Experimental mice

In this appendix, mice were labelled according to their cage number, listed below:

JL1041	Donor RAG2$^{-/-}$ mouse (in which *P. chabaudi* paraites were expanded prior to infection of the challenge mice).
JL1031	Immunized with saline in the presence of adjuvant
JL1032	Immunized with $MSP1_{p21}$
JL1033	Immunized with the conserved CIR peptide
JL1034	Immunized with the CIR sub-family U peptide
JL1035	Immunized with the recombinant CIR PCHAS_040110
JL1036	Immunized with the CIR sub-family A-F peptide
JL1037	Immunized with the recombinant CIR PCHAS_000100
JL1038	Immunized with the CIR cocktail

Appendix 6.3 Peptide immunization cross-reactivity

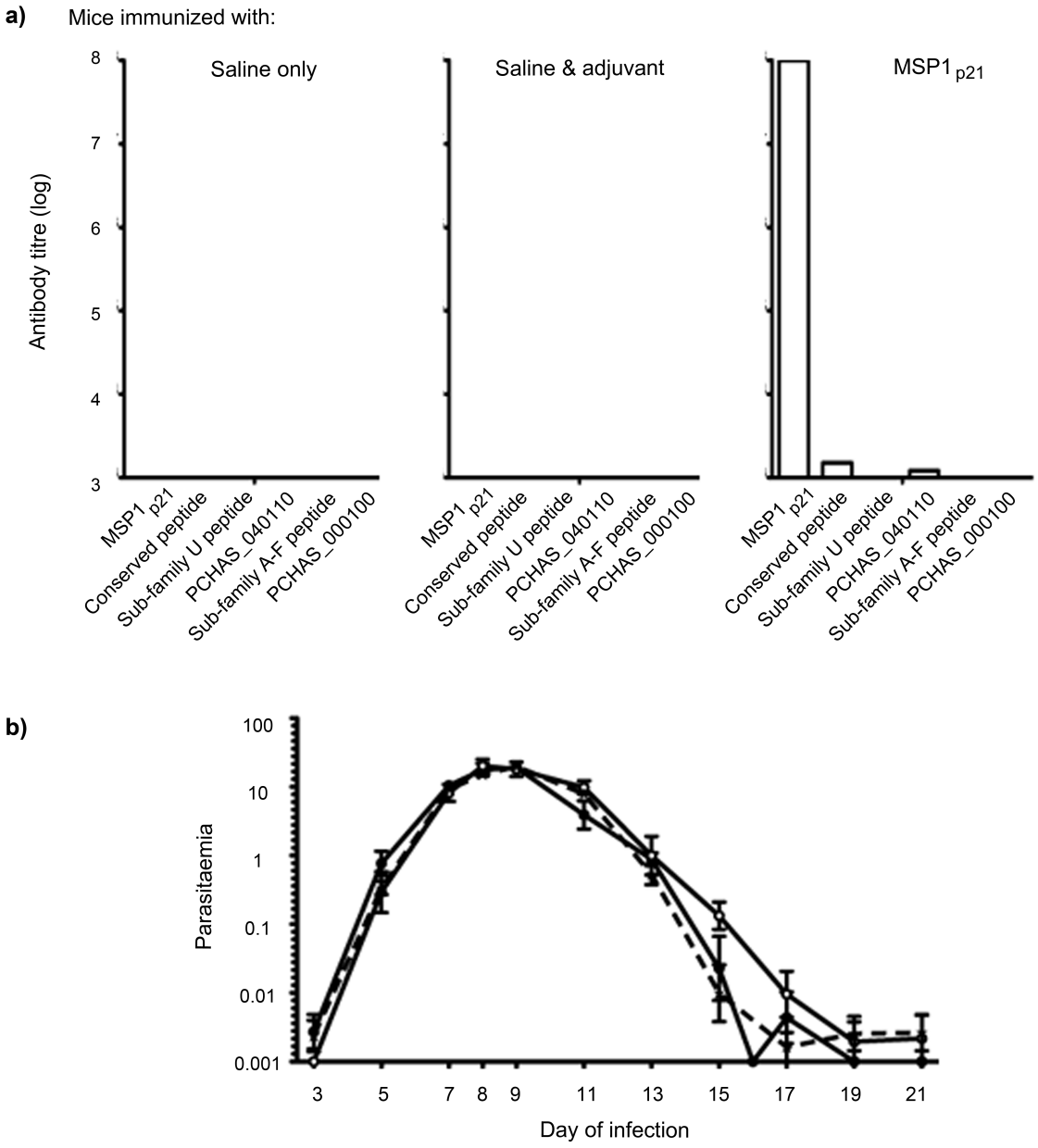
ELISAs were performed using the un-conjugated sub-family U and A-F peptides, to determine whether mice immunized with CIR peptides made cross-reactive Abs recognizing other peptides, as described in chapter 6. Mice immunized with a CIR peptide made Abs recognizing that peptide, but not the other CIR peptide(s).



The reactivity of sera from mice immunized with the conserved CIR peptide, sub-family specific peptides or the CIR cocktail is colour coded as follows, with filled symbols: conserved CIR peptide, purple; CIR cocktail, pink; CIR sub-family U peptide, orange; CIR sub-family A-F peptide, light blue.

In addition, pre-immune sera from the same mice prior to immunization are plotted according to the same colour code, using hollow symbols. The mean of three technical replicates is plotted, as a two-fold titration from a starting sera dilution of 1/100. Error bars represent the standard error of the mean.

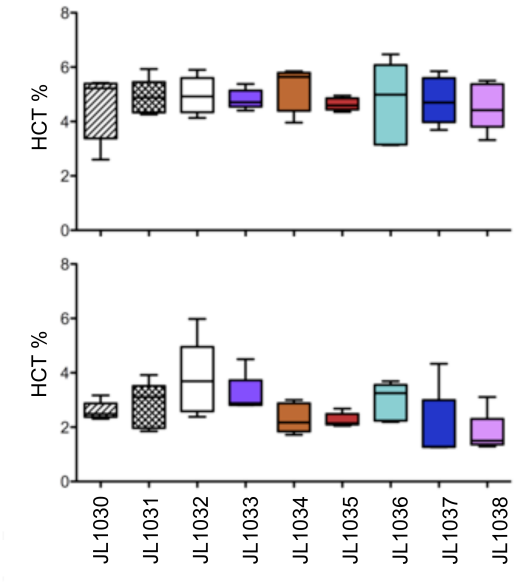
Appendix 6.4 *P. chabaudi* challenge of $MSP1_{p21}$ immunized mice



Titres of Abs recognizing $MSP1_{p21}$ and CIR proteins and peptides were measured in the sera of mice immunized with saline alone or in the presence of adjuvant, and $MSP1_{p21}$ in the presence of adjuvant, **a)**. Titre was defined as the dilution of immunized serum at which it reached the OD405 nm of serum from the same mouse prior to immunization, calculated using the mean of two technical replicates. Error bars represent the standard error of the mean.

The geometric mean parasitaemia for these three groups of 5 mice are shown, **b)**. Error bars represent the standard error of the mean.

Appendix 6.5 Anaemia during *P. chabaudi* infection



Blood samples were taken from challenge mice on day 3 and day 8 of *P. chabaudi* infection. These were used to calculate the haematocrit (HCT%) of the blood for all mice, which is a measure of anaemia. The mean HCT% and 95 percentiles are plotted for each group of 5 mice. The groups are labelled according to their cage numbers.

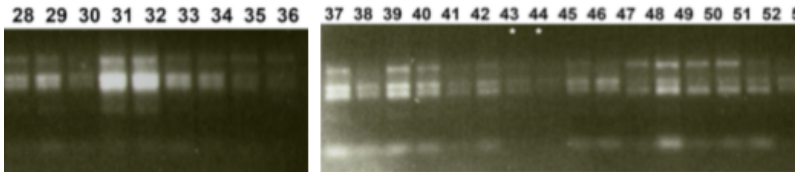
Appendix 6.6 RNA integrity

Appendix 6.6.1 Spectroscopy. (M refers to individual mice).

Sample	Tube	Concentration (ng/ul)	500ng	A260/A280	Position in 96 well plate
JL1031 challenge	M1	225.59	2.22	1.92	F5
	M2	143.36	3.49	1.86	F6
	M5	111.19	4.50	1.83	F7
JL1032 challenge	M3	197.38	2.53	1.87	F8
	M4	214.41	2.33	1.85	F9
	M5	147.41	3.39	1.88	F10
JL1033 challenge	M2	196.92	2.54	1.93	F11
	M3	172.33	2.90	1.87	F12
	M4	163.15	3.06	1.92	G1
JL1034 challenge	M1	222.11	2.25	1.92	G2
	M3	109.61	4.56	1.96	G3
	M4	78.32	6.38	1.84	G4
JL1035 challenge	M1	226.15	2.21	1.89	G5
	M4	108.19	4.62	1.82	G6
	M5	603.73	0.83	1.74	G7
JL1036 challenge	M1	230.97	2.16	1.89	G8
	M3	178.79	2.80	1.92	G9
	M5	95.48	5.24	1.82	G10
JL1037 challenge	M1	305.13	1.64	1.99	G11
	M2	231.43	2.16	1.92	G12
	M3	235.86	2.12	1.93	H1
JL1038 challenge	M1	95.48	5.24	1.81	H2
	M2	127.39	3.92	1.86	H3
	M3	145.02	3.45	1.90	H4
Donor JL1041-2	1	561.77	0.89	1.69	H5
	2	601.67	0.83	1.82	H6

Appendix 6.6.2 Gel electrophoresis of 500ng RNA

RNA was extracted from these samples at the same time as those in Appendix 3.7.3 (the numbering continues from these).



28) JL1031 M1
29) JL1031 M2
30) JL1031 M5
31) JL1032 M3
32) JL1032 M4
33) JL1032 M5
34) JL1033 M2
35) JL1033 M3
36) JL1033 M4
37) JL1034 M1
38) JL1034 M3
39) JL1034 M4
40) JL1035 M1
41) JL1035 M4
42) JL1035 M5
43) JL1036 M1
44) JL1036 M3
45) JL1036 M5
46) JL1037 M1
47) JL1037 M2
48) JL1037 M3
49) JL1038 M1
50) JL1038 M2
51) JL1038 M3
52) JL1041 M1

Appendix 6.7 RT-qPCR analysis

Analyzed as described in chapter 3 and Appendix 3.5.1.

Mouse:	U1			U4			C1			E1		
	Mean ratio	SD	N	Mean ratio	SD	N	Mean ratio	SD	N	Mean ratio	SD	N
JL1031-1	4.10	0.22	3.00	2.33	0.39	3.00	1.05	0.22	3.00	1.65	0.27	3.00
JL1031-2	11.63	0.93	3.00	0.41	0.22	2.00	4.45	0.29	3.00	1.46	0.24	3.00
JL1031-5	1.92	0.15	3.00	0.19	0.29	3.00	0.88	0.33	3.00	1.63	0.00	3.00
JL1032-3	1.19	0.29	3.00	0.15	0.19	3.00	0.73	0.22	3.00	0.44	0.07	3.00
JL1032-4	3.44	0.16	3.00	0.58	0.27	3.00	1.37	0.04	3.00	0.91	0.12	3.00
JL1032-5	3.39	0.17	3.00	0.22	0.13	3.00	1.45	0.09	3.00	0.97	0.18	3.00
JL1033-2	0.35	0.06	3.00	0.04	0.18	3.00	0.14	0.22	3.00	0.16	0.22	3.00
JL1033-3	1.39	0.29	3.00	0.15	0.14	3.00	0.53	0.59	3.00	0.41	0.39	3.00
JL1033-4	0.91	0.17	3.00	0.12	0.22	3.00	1.05	2.06	3.00	0.39	0.28	3.00
JL1034-1	0.76	0.05	3.00	0.10	0.13	3.00	1.50	0.00	3.00	6.79	0.54	3.00
JL1034-3	1.67	0.14	3.00	0.44	0.82	2.00	4.26	0.00	3.00	15.99	0.00	2.00
JL1034-4	3.45	0.47	3.00	1.40	0.27	3.00	0.70	0.00	3.00	2.64	0.00	3.00
JL1035-1	2.29	0.09	3.00	0.32	0.01	3.00	0.44	0.00	3.00	10.16	0.20	3.00
JL1035-4	2.53	0.08	3.00	0.25	0.06	3.00	1.26	0.00	3.00	4.73	0.00	2.00
JL1035-5	0.59	0.25	3.00	0.13	0.11	3.00	6.63	2.66	3.00	28.18	1.93	3.00
JL1036-1	2.81	0.38	3.00	2.05	1.30	3.00	0.51	0.00	3.00	1.92	0.00	3.00
JL1036-3	4.10	0.00	3.00	1.62	0.03	2.00	0.69	0.07	3.00	1.56	0.43	2.00
JL1036-5	11.63	0.00	3.00	4.53	0.00	3.00	42.70	1.56	3.00	262.96	1.61	3.00
JL1037-1	1.92	0.02	3.00	0.75	0.06	3.00	0.88	0.11	3.00	13.02	3.53	3.00
JL1037-2	1.19	0.00	3.00	0.46	0.00	3.00	4.25	3.99	3.00	115.67	4.95	3.00
JL1037-3	3.44	0.00	3.00	1.44	0.63	3.00	0.37	0.19	3.00	3.99	0.64	2.00
JL1038-1	3.39	0.00	3.00	3.57	1.01	3.00	0.82	0.81	3.00	0.93	0.19	2.00
JL1038-2	1.39	0.00	3.00	0.54	0.00	3.00	20.66	4.88	3.00	3.29	0.75	3.00
JL1038-3	1.08	0.47	3.00	0.14	0.59	3.00	0.59	0.30	3.00	0.70	0.23	3.00